

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 18:12:45 ; Search time 7189.25 Seconds  
(without alignments)  
17088.240 Million cell updates/sec

Title: US-09-935-290-1  
Perfect score: 3003  
Sequence: 1 ttccggaccaggctgtcg.....cttgagatcttcccagcttt 3003

Scoring table: IDENTITY\_NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sv.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3001	99.9	3003	6	AX420453	AX420453 Sequence
2	2958	98.5	3210	6	AX451419	AX451419 Sequence
3	2755.4	91.8	4918	9	HSM804404	AL833093 Homo sapi
4	2752.2	91.6	3490	9	BC030783	BC030783 Homo sapi
5	2752.2	91.6	6387	9	HSM804372	AL833061 Homo sapi
6	2485.4	82.8	2487	6	AX431391	AX431391 Sequence
7	2484	82.7	2484	6	AX420455	AX420455 Sequence
8	2241.8	74.7	5689	9	HSM803772	AL832464 Homo sapi
9	2131	71.0	5674	9	AB045780	AB045780 Homo sapi
10	1917.6	63.9	3857	10	BC019201	BC019201 Mus muscu
11	1903.8	63.4	6634	10	MUSG3PAT	M77003 Mouse glyce
12	1896.2	63.1	2646	6	AX401972	AX401972 Sequence
13	1896.2	63.1	2646	10	AF021348	AF021348 Rattus no
14	1807.2	60.2	2519	10	RNU36771	U36771 Rattus norv
15	267.4	8.9	196657	9	AL391986	AL391986 Human DNA
16	246.4	8.2	271	6	BD029418	BD029418 Sequence
17	223.6	7.4	255430	2	AC099387	AC099387 Rattus no
18	218.2	7.3	197741	2	AC110206	AC110206 Mus muscu
19	203.4	6.8	2000	6	AX431393	AX431393 Sequence
20	173.8	5.8	206914	2	AC104911	AC104911 Mus muscu
21	142.6	4.7	4201	4	AF469047	AF469047 Bos tauru
22	131.6	4.4	147242	2	AC135152	AC135152 Rattus no
23	123.2	4.1	147242	2	AC135152	AC135152 Rattus no
24	109	3.6	2683	6	AX710670	AX710670 Sequence
25	109	3.6	2755	6	AX451420	AX451420 Sequence
26	107.8	3.6	2382	6	AX710668	AX710668 Sequence
27	107.8	3.6	2388	6	AX710666	AX710666 Sequence
28	107.8	3.6	2665	6	AX710664	AX710664 Sequence
29	103.8	3.5	3450	3	AF145635	AF145635 Drosophil
30	103.8	3.5	76299	2	AC019585	AC019585 Drosophil
31	103.8	3.5	172363	3	AC009388	AC009388 Drosophil
32	103.8	3.5	206190	3	AE003763	AE003763 Drosophil
33	95.4	3.2	539	11	G98364	G98364 S210P6250FD
34	85.6	2.9	38145	3	U64847	U64847 Caenorhabdi
35	75	2.5	40532	3	CBRG36D18	CBRG36D18
36	64.8	2.2	301995	1	AE016779	AE016779 Pseudomon
37	61.4	2.0	10744	1	AE013659	AE013659 Yersinia
38	61.4	2.0	214050	1	AJ414142	AJ414142 Yersinia
39	59.2	2.0	300029	1	AE012554	AE012554 Xylella f
40	58	1.9	310029	1	AE016861	AE016861 Pseudomon
41	56.6	1.9	299938	1	AE016800	AE016800 Vibrio vu
42	56	1.9	11392	1	AE003940	AE003940 Xylella f
43	55	1.8	10802	1	AE015425	AE015425 Shigella
44	55	1.8	292497	1	AE016989	AE016989 Shigella
45	54.6	1.8	221078	2	AC133100	AC133100 Mus muscu

ALIGNMENTS

RESULT 1	AX420453	AX420453	3003 bp	DNA	linear	PAT 18-JUN-2002
AX420453	Sequence 1 from Patent WO0216592.					
LOCUS	AX420453					
DEFINITION	Sequence 1 from Patent WO0216592.					
ACCESSION	AX420453					
VERSION	AX420453.1	GI:21524602				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Kapeller-Libermann, R. and Logan, T.J.					
TITLE	Atcr-1, a human acyltransferase and uses thereof					
JOURNAL	Patent: WO 0216592-A 1 28-FEB-2002;					

Millennium Pharmaceuticals, Inc. (US)	
FEATURES	Location/Qualifiers
source	1..3003
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
CDS	341..2827
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAD35486.1"
	/db_xref="GI:21524603"
	/translation="WDESLTLGTDVSLPHSSEYVGRCKHTSEWEGCGRPRTVF RSATLWKESLMRKPFPVGRCCYSCDTPQSMDFNPSIPSLGLRNVINYNETHTRR GWLARLSYVLFIQEDVHKGFMATNTENVLSNRVQEAIAEVAELNPDGSAOQS KANVKKKAKRILOEMVATVSPAMIRLTGWLLKLPNSFPFNIQIHKGLEWVKAA ETNLPILFVPHRSHDYLLLTFLFCHNIKAPIASGNLNIPIFTSLHKLGGEFI RRLEDETPGRKDVYRLQHQFLEIFLEGTRSRSGKTS CARAGLLSV VVDLTNNVPDILIIIPVGISYDRIIEGHNGEOLKPKKNESLWSVARGVIRMLRN YGCVRVDFAPFSLKEYLESQSPVLSALSLEQALLPALPRSPDADEGRDTSIN ESRNATDESURRLIANLAHEILFTASKSAIMSTHIVACLILYRHRQGDIDLSLVED FVYKKEVLARDPDLFGNSGSDVVMHAIQLLGNCSVTIHTSRNDEFFIPTSTVPSV FELNFSYNGVLHVFIMEAIIACSLYAVLNKRGGLGPTSPNLIISQELVRKAASLCY LLNSGNTSLPQCTFYQVCHETVKFIOYGLITVAEHDDQEDISPLAEQOQWKKLPE PLSWRDEDEDDFGEQRQDCYLKVSQSKHQFPITFLQRLGLPLEAYSSAIIKVH NFGSPVPEPYLQKLHKYLITRERNVAVAESATYCLVKNVAMFKDIGVFKEIKQK RVSVLELSSTFLPQCNKQKLEVLISFVL"
BASE COUNT	826 a 683 c 698 g 794 t 2 others
ORIGIN	
	Query Match 99.9%; Score 3001; DB 6; Length 3003;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 3003; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 TTGGGACACAGGCTGCTGGCGGGGACTCTTTCTGAGGTTACTGTGGAGCACCCAAAGTC 60
Db	1 TTGGGACCAAGGCTGCTGGCGGGGACTCTTTCTGAGGTTACTGTGGAGCACCCAAAGTC 60
Qy	61 TGTCAGGCTCTGGCGGTGCAACAGGACCCAGAGGAACAGACCTTGTCTATTTCACCCA 120
Db	61 TGTCAGGCTCTGGCGGTGCAACAGGACCCAGAGGAACAGACCTTGTCTATTTCACCCA 120
Qy	121 CAGCTGGGACTGTCTTCCAGAGTCTCCATCAGCTTTTGTCTAATCGACTGATTTGGAAT 180
Db	121 CAGCTGGGACTGTCTTCCAGAGTCTCCATCAGCTTTTGTCTAATCGACTGATTTGGAAT 180
Qy	181 AATTCTCAACACCAACCAAGTCAAGATACAGGACAGCGGCTCCCTGTGTATGGA 240
Db	181 AATTCTCAACACCAACCAAGTCAAGATACAGGACAGCGGCTCCCTGTGTATGGA 240
Qy	241 CATTCTGCACCCGAACTGATAGCTGAGTCTCTGAAGTTTATGTTATGAACAGAGAAC 300
Db	241 CATTCTGCACCCGAACTGATAGCTGAGTCTCTGAAGTTTATGTTATGAACAGAGAAC 300
Qy	301 TTTTCATCCCGACCATGATTTGGGAATTACACTTTGTGACATGGATGAATCTGCACGTAC 360
Db	301 TTTTCATCCCGACCATGATTTGGGAATTACACTTTGTGACATGGATGAATCTGCACGTAC 360
Qy	361 CCTTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCAGATACAGATGTTGGTCGATG 420
Db	361 CCTTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCAGATACAGATGTTGGTCGATG 420
Qy	421 TAAGCACACAAGTACAGAAATGGGTGAGTGGCTTTTAGACCCACCGCTTTCAGATCTGC 480
Db	421 TAAGCACACAAGTACAGAAATGGGTGAGTGGCTTTTAGACCCACCGCTTTCAGATCTGC 480
Qy	481 AACCTTTAAATGGAAGAAAGCCCTAATAGTTCGGAAGGCCCAATTTGTTGGAAGATGTTG 540
Db	481 AACCTTTAAATGGAAGAAAGCCCTAATAGTTCGGAAGGCCCAATTTGTTGGAAGATGTTG 540
Qy	541 TTACTCTGCATCCCGACAGCTGGGACAAATTTTTCACCCCGATATCCCGTCTTTGGG 600
Db	541 TTACTCTGCATCCCGACAGCTGGGACAAATTTTTCACCCCGATATCCCGTCTTTGGG 600

Qy	601	TTTTCGGGAATGTTATTATATCAATGAACTCACACAAGCACCGCGGATGGCTTGC	660
Db	601	TTTTCGGGAATGTTATTATATCAATGAACTCACACAAGCACCGCGGATGGCTTGC	660
Qy	661	ACGCCCTTTCTTACGCTTCTTTTATTTCAAGAGCGAGATGTGCATAAGGGCATGTTTGGCCAC	720
Db	661	ACGCCCTTTCTTACGCTTCTTTTATTTCAAGAGCGAGATGTGCATAAGGGCATGTTTGGCCAC	720
Qy	721	CAATGTGACTGAAAAATGTCTGAAACAGCAGCTAGAGTACAAGAGGCAATTCGACGAAGTGGC	780
Db	721	CAATGTGACTGAAAAATGTCTGAAACAGCAGCTAGAGTACAAGAGGCAATTCGACGAAGTGGC	780
Qy	781	TGCTGAAATTAACCCCTGATGTGTTCTGCCAGCAGCAATCAAAAGCCGTTTAACAAAGTGAA	840
Db	781	TGCTGAAATTAACCCCTGATGTGTTCTGCCAGCAGCAATCAAAAGCCGTTTAACAAAGTGAA	840
Qy	841	AAAGAAAGCTAAAGGATTTCTTCAAGAAATGSGTTCGCCACTGTCTCACCGGCAATGATCAG	900
Db	841	AAAGAAAGCTAAAGGATTTCTTCAAGAAATGSGTTCGCCACTGTCTCACCGGCAATGATCAG	900
Qy	901	ACTGACTGGGTGGGTGCTGCTTAAACACTGTTTCAACAGCTTCTTTTGGAAACATTTCAAATTC	960
Db	901	ACTGACTGGGTGGGTGCTGCTTAAACACTGTTTCAACAGCTTCTTTTGGAAACATTTCAAATTC	960
Qy	961	CAAGGTCACACTTGAGATGGTTAAAGCTGCAACTGTGACGAAATTTGCGGCTTCTGTGTTCT	1020
Db	961	CAAGGTCACACTTGAGATGGTTAAAGCTGCAACTGTGACGAAATTTGCGGCTTCTGTGTTCT	1020
Qy	1021	ACCAGTTCATAGATCCCATATTGACTATCTGCTGCTCCTCCTCATCTTCTTTCGCCATAA	1080
Db	1021	ACCAGTTCATAGATCCCATATTGACTATCTGCTGCTCCTCCTCATCTTCTTTCGCCATAA	1080
Qy	1081	CATCAAGACCATACATTTGCTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTT	1140
Db	1081	CATCAAGACCATACATTTGCTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTT	1140
Qy	1141	GATCCATAAGCTTGGGGCTTCTTCATACGACGAAGGCTCGATGAAACACACAGATGGACG	1200
Db	1141	GATCCATAAGCTTGGGGCTTCTTCATACGACGAAGGCTCGATGAAACACACAGATGGACG	1200
Qy	1201	GAAAGATGTTCTATATAGAGCTTTGCTCCATGCGCATATAGTTGAAATTACTTCGACAGCA	1260
Db	1201	GAAAGATGTTCTATATAGAGCTTTGCTCCATGCGCATATAGTTGAAATTACTTCGACAGCA	1260
Qy	1261	GCAATTTCTGGAGATCTTCTCGGAAGGCACACGTTCTTAGGAGTGGAAAAACCTCTTGTC	1320
Db	1261	GCAATTTCTGGAGATCTTCTCGGAAGGCACACGTTCTTAGGAGTGGAAAAACCTCTTGTC	1320
Qy	1321	TCGGGCAGGACTTTTGTGAGTGTGGTAGATCTCTGTCTACCAATGTCATCCAGACAT	1380
Db	1321	TCGGGCAGGACTTTTGTGAGTGTGGTAGATCTCTGTCTACCAATGTCATCCAGACAT	1380
Qy	1381	CTTGATATACTGTTTGAATCTCTTATGATCGCAATTATCGAAGGTCACTTACAATGTTGA	1440
Db	1381	CTTGATATACTGTTTGAATCTCTTATGATCGCAATTATCGAAGGTCACTTACAATGTTGA	1440
Qy	1441	ACAACTGGGCAAAACCTAAGAAAGATGAGAGCTCTGGAGTGTAGCAAGAGGTGTTATTAG	1500
Db	1441	ACAACTGGGCAAAACCTAAGAAAGATGAGAGCTCTGGAGTGTAGCAAGAGGTGTTATTAG	1500
Qy	1501	AATGTTACGAAAAAACTATGGTTGTGCCAGTGGATTTTTCACAGCCATTTTCTTTAAA	1560
Db	1501	AATGTTACGAAAAAACTATGGTTGTGCCAGTGGATTTTTCACAGCCATTTTCTTTAAA	1560
Qy	1561	GGAATATTTAGAAGCCAAAAGTCAGAAACCGGTGTCTGCTCTACTTTCCCTGGAGCAAGC	1620
Db	1561	GGAATATTTAGAAGCCAAAAGTCAGAAACCGGTGTCTGCTCTACTTTCCCTGGAGCAAGC	1620
Qy	1621	GTTGTTTACCGCTATCTTCTTCAAGCCAGGTGATGCTGCTGATCAAGGTAGAGACAC	1680
Db	1621	GTTGTTTACCGCTATCTTCTTCAAGCCAGGTGATGCTGCTGATCAAGGTAGAGACAC	1680
Qy	1681	GTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAAGGAGTTGATTGCAAA	1740

Db 1681 GTCCATTATAGTCCAGAAATGCACAGATGAATCCCTTACGAAGAGGTGATTCGCAAA 1740  
Qy 1741 TCTGGCTGAGCATATTTCTATTCACTGCTAGCAAGTCCCTGTGCGCATATATGTTCACACACAT 1800  
Db 1741 TCTGGCTGAGCATATTTCTATTCACTGCTAGCAAGTCCCTGTGCGCATATATGTTCACACACAT 1800  
Qy 1801 TGTGGCTTGGCTGCTCTCTACAGACAGAGGAGGAAATGATCTCTCAATTTGGTGA 1860  
Db 1801 TGTGGCTTGGCTGCTCTCTACAGACAGAGGAGGAAATGATCTCTCAATTTGGTGA 1860  
Qy 1861 AGACTTCTTTGTGATGAAGAGAGAGTCTGTGCTCGTATTTTGACCTGGGGTCTCAGG 1920  
Db 1861 AGACTTCTTTGTGATGAAGAGAGAGTCTGTGCTCGTATTTTGACCTGGGGTCTCAGG 1920  
Qy 1921 AAATTCAGAAATGTAGTAATCATGATGCCATACAGCTGCTGGGAAATTTGTGCACAAATCAC 1980  
Db 1921 AAATTCAGAAATGTAGTAATCATGATGCCATACAGCTGCTGGGAAATTTGTGCACAAATCAC 1980  
Qy 1981 CCACACTAGCAGGAACGATGAGTTTTTATCAACCCCGAGCAAACTGTCCCATCATGCTTT 2040  
Db 1981 CCACACTAGCAGGAACGATGAGTTTTTATCAACCCCGAGCAAACTGTCCCATCATGCTTT 2040  
Qy 2041 CGAACTCAACTTCTACAGCAATGGGTACTTCTATCTTTATCATGAGGAGCCATCATAGC 2100  
Db 2041 CGAACTCAACTTCTACAGCAATGGGTACTTCTATCTTTATCATGAGGAGCCATCATAGC 2100  
Qy 2101 TTGACGCTTTATGCAATGAGGAGTCTGCAACAGAGGAGTCTGGGGTCCCATCTAGCACCCACC 2160  
Db 2101 TTGACGCTTTATGCAATGAGGAGTCTGCAACAGAGGAGTCTGGGGTCCCATCTAGCACCCACC 2160  
Qy 2161 TAACTGTATGACCCAGGAGCAGTCTGTGCGGAAGCGCCGACCTGTGCTACCTTCTCTC 2220  
Db 2161 TAACTGTATGACCCAGGAGCAGTCTGTGCGGAAGCGCCGACCTGTGCTACCTTCTCTC 2220  
Qy 2221 CAATCAAGGACCATCTCACTGCTTCCAGACATTTTACCAGTCTGCCATGAACAGT 2280  
Db 2221 CAATCAAGGACCATCTCACTGCTTCCAGACATTTTACCAGTCTGCCATGAACAGT 2280  
Qy 2281 AGGAAGTTTATCCAGTATGGGATTTTACAGTGGCAGACGATGATCCAGGAAGATAT 2340  
Db 2281 AGGAAGTTTATCCAGTATGGGATTTTACAGTGGCAGACGATGATCCAGGAAGATAT 2340  
Qy 2341 CAGTCTAGTCTGTGAGCAGCAGTGGGACAAAGAGCTTCCAGAACCTTTGTCTTGAG 2400  
Db 2341 CAGTCTAGTCTGTGAGCAGCAGTGGGACAAAGAGCTTCCAGAACCTTTGTCTTGAG 2400  
Qy 2401 AAGTATGAAGAGATGAAGACAGTACTTTGGGGAGGAACAGCGAGATTTGCTACTGAA 2460  
Db 2401 AAGTATGAAGAGATGAAGACAGTACTTTGGGGAGGAACAGCGAGATTTGCTACTGAA 2460  
Qy 2461 GGTGAGCCAAATCCAGGAGCACAGCAGTTTATCACCTTCTTACAGAGACTCTTTGGGCC 2520  
Db 2461 GGTGAGCCAAATCCAGGAGCACAGCAGTTTATCACCTTCTTACAGAGACTCTTTGGGCC 2520  
Qy 2521 TTTGCTGAGGCTTACAGCTCTGTGCGCATCTTTGTTTCACAACTTTCAGTGGTCTGTTC 2580  
Db 2521 TTTGCTGAGGCTTACAGCTCTGTGCGCATCTTTGTTTCACAACTTTCAGTGGTCTGTTC 2580  
Qy 2581 AGAACCCTGATATCTGCAAAAAGTTGCAAAATACCTAATAACAGAAAGAAATGT 2640  
Db 2581 AGAACCCTGATATCTGCAAAAAGTTGCAAAATACCTAATAACAGAAAGAAATGT 2640  
Qy 2641 TGCATATATGCTGAGAGTCCACATATTTGTTGAGAGATGCTGTGAAAATGTTTTAA 2700  
Db 2641 TGCATATATGCTGAGAGTCCACATATTTGTTGAGAGATGCTGTGAAAATGTTTTAA 2700  
Qy 2701 GGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTGTCTGTTTTAGAACTGAGCAG 2760  
Db 2701 GGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTGTCTGTTTTAGAACTGAGCAG 2760  
Qy 2761 CACTTTTCTACTCAATCAACCCGACAAAATCTCTAGAATATATTTGAGTTTGTGGT 2820

Db 2761 CACTTTTCTACTCAATCAATGCAACCCGACAAAATCTCTAGAATATATTTCTGAGTTTGTGGT 2820  
Qy 2821 GCTGTAGTAACTGTGTGGCACTGCTGGCAAAATGAAGGTCAAGATGAGTTCTTTGTAGG 2880  
Db 2821 GCTGTAGTAACTGTGTGGCACTGCTGGCAAAATGAAGGTCAAGATGAGTTCTTTGTAGG 2880  
Qy 2881 TACCAGCTTCTGGCTCAAGAGTTTGAAGTCCCTTTCGAGGGGTGAGCCCTGCTGTC 2940  
Db 2881 TACCAGCTTCTGGCTCAAGAGTTTGAAGTCCCTTTCGAGGGGTGAGCCCTGCTGTC 2940  
Qy 2941 CGAAGTGTATCTCTGGAGACAGTCTTCTNCTCCATGATCTGAGATCTTCCCGC 3000  
Db 2941 CGAAGTGTATCTCTGGAGACAGTCTTCTNCTCCATGATCTGAGATCTTCCCGC 3000  
Qy 3001 TTT 3003  
Db 3001 TTT 3003  
  
RESULT 2  
AX451419  
LOCUS AX451419 3210 bp DNA linear PAT 03-JUL-2002  
DEFINITION Sequence 26 from Patent WO0226950.  
ACCESSION AX451419  
VERSION AX451419.1 GI:21698428  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Lal, P.G., Tang, Y.T., Yue, H., Burford, N., Gandhi, A.R., Warren, B.A.,  
Yao, M.G., Tribouley, C.M., Baughn, M.R., Lee, E.A., Hafalia, A.J.,  
Lu, Y., Griffin, J.A., Sanjanwala, M.S. and Ding, L.  
TITLE Transferrases  
JOURNAL Patent: WO 0226950-A 26 04-APR-2002;  
Incyte Genomics, Inc. (US)  
FEATURES  
Location/Qualifiers  
source 1..3210  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 1517230CB1"  
BASE COUNT 896 a 721 c 735 g 858 t  
ORIGIN  
Query Match 98.5%; Score 2958; DB 6; Length 3210;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2983; Conservative 0; Mismatches 7; Indels 2; Gaps 2;  
Qy 12 GCTGCTGGGGGACCTTTTCTGAGGTTACTGTGGAGCACCCAAAGTCTGTCAGCTCT 71  
Db 9 GCTGCTGGGGGACCTTTTCTGAGGTTACTGTGGAGCACCCAAAGTCTGTCAGCTCT 68  
Qy 72 GGCCGTGCAACAGCGCACCCAGAGAAACAGACCTTGTATTATCCACAGCCTGGGAC 131  
Db 69 GGCCGTGCAACAGCGCACCCAGAGAAACAGACCTTGTATTATCCACAGCCTGGGAC 128  
Qy 132 TGTCTTCTCCAGAGTCTCCATCAGCTTTGTCTTAATGATGAGTGAATAATTCCTCAA 191  
Db 129 TGTCTTCTCCAGAGTCTCCATCAGCTTTGTCTTAATGATGAGTGAATAATTCCTCAA 188  
Qy 192 CACCACCAAGTCAAGGATACAGGAGCGGGCTCCCTGTTGATGAGACATTCGAC 251  
Db 189 CACCACCAAGTCAAGGATACAGGAGCGGGCTCCCTGTTGATGAGACATTCGAC 248  
Qy 252 CGAAACTGATAGTCTGAGTCTTGAAGTATTATGTTATGAAACAGAAAGAACTTTCATCC 311  
Db 249 CGAAACTGATAGTCTGAGTCTTGAAGTATTATGTTATGAAACAGAAAGAACTTTCATCC 308  
Qy 312 CACATGATTTGGGAAATACACTTTTGTGACATGAGATGAATCTGCACTGACCTTGGTACAA 371  
Db 309 CACATGATTTGGGAAATACACTTTTGTGACATGAGATGAATCTGCACTGACCTTGGTACAA 368

372 TAGATGTTTCTTATCTGCCACATTCATCAGAAATACAGTGTGGTTCGATGTAAAGCACAA 431  
Db |||||  
369 TAGATGTTTCTTATCTGCCACATTCATCAGAAATACAGTGTGGTTCGATGTAAAGCACAA 428  
Qy |||||  
432 GTGAGGAATGGGTGAGTGTGGCTTTAGACCCACCGTCTTCAGATCTGCAACTTTAAAT 491  
Db |||||  
429 GTGAGGAATGGGTGAGTGTGGCTTTAGACCCACCATCTTCAGATCTGCAACTTTAAAT 488  
Qy |||||  
492 GGAAGGAAGCCTAATGAGTCGGAAGAGGCCATTTGTTGAAGATGTTGTTACTCTGCA 551  
Db |||||  
489 GGAAGGAAGCCTAATGAGTCGGAAGAGGCCATTTGTTGAAGATGTTGTTACTCTGCA 548  
Qy |||||  
552 CTCCCGAGAGCTGGGCAAAATTTTCAACCCCGAGTATCCCGTCTTTGGGTTTCGGAATG 611  
Db |||||  
549 CTCCCGAGAGCTGGGCAAAATTTTCAACCCCGAGTATCCCGTCTTTGGGTTTCGGAATG 608  
Qy |||||  
612 TTATTATATCAATGAAATCAACAAGACACCGCGGATGGCTTGCAAGACGCTTTCTT 671  
Db |||||  
609 TTATTATATCAATGAAATCAACAAGACACCGCGGATGGCTTGCAAGACGCTTTCTT 668  
Qy |||||  
672 ACCTTCTTTTATCAAGAGGAGATGTGCATTAAGGCAATGTTGCCACCAATGTCACTG 731  
Db |||||  
669 ACCTTCTTTTATCAAGAGGAGATGTGCATTAAGGCAATGTTGCCACCAATGTCACTG 728  
Qy |||||  
732 AAAATGTGCTGAACAGCAGTAGAGTACAAGAGCAATTTGCAGAGTGGCTGCTGAATTA 791  
Db |||||  
729 AAAATGTGCTGAACAGCAGTAGAGTACAAGAGCAATTTGCAGAGTGGCTGCTGAATTA 788  
Qy |||||  
792 ACCCTGATGGTTCTGCCAGCAGCAATCAAAAGCCGTTAACAAAGTGAAGAAAGACTA 851  
Db |||||  
789 ACCCTGATGGTTCTGCCAGCAGCAATCAAAAGCCGTTAACAAAGTGAAGAAAGACTA 848  
Qy |||||  
852 AAAGGATCTTCAAGAAATGGTTGCCACTGTCTCACCGCAATGATCAGACTGCGGT 911  
Db |||||  
849 AAAGGATCTTCAAGAAATGGTTGCCACTGTCTCACCGCAATGATCAGACTGCGGT 908  
Qy |||||  
912 GGGTGTCTGCTAAATCTGTTCAACAGCTTTCTTTGGAAACATTCAAATTCACAAAGTCAAC 971  
Db |||||  
909 GGGTGTCTGCTAAATCTGTTCAACAGCTTTCTTTGGAAACATTCAAATTCACAAAGTCAAC 968  
Qy |||||  
972 TTGAGATGGTTAAAGCTGCAACTGAGACGAATTTGCGCTTCTGTTTCTACAGTTTATA 1031  
Db |||||  
969 TTGAGATGGTTAAAGCTGCAACTGAGACGAATTTGCGCTTCTGTTTCTACAGTTTATA 1028  
Qy |||||  
1032 GATCCCATATTTGATATCTGCTGCTCATTTTCAATCTCTTCTGCCATAACATCAAGCAC 1091  
Db |||||  
1029 GATCCCATATTTGATATCTGCTGCTCATTTTCAATCTCTTCTGCCATAACATCAAGCAC 1088  
Qy |||||  
1092 CATACATGCTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTGATCCATAAGC 1151  
Db |||||  
1089 CATACATGCTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTGATCCATAAGC 1148  
Qy |||||  
1152 TTGGGGCTTCTTCATACGACGAAGGCTCGATGAACACACAGATGGACGGAAGATGTTTC 1211  
Db |||||  
1149 TTGGGGCTTCTTCATACGACGAAGGCTCGATGAACACACAGATGGACGGAAGATGTTTC 1208  
Qy |||||  
1212 TCTATAGAGCTTGTCTCATGGGCATATAGTTGAAATTAATTCGACAGAGCAATTCCTGG 1271  
Db |||||  
1209 TCTATAGAGCTTGTCTCATGGGCATATAGTTGAAATTAATTCGACAGAGCAATTCCTGG 1268  
Qy |||||  
1272 AGATCTTCGGAAGGCACACGTTCTAGGAGTGGAAACCTCTTGTGCTCGGGCAGGAC 1331  
Db |||||  
1269 AGATCTTCGGAAGGCACACGTTCTAGGAGTGGAAACCTCTTGTGCTCGGGCAGGAC 1328  
Qy |||||  
1332 TTTTGTGAGTGTGGTAGATACTCTGTCTACCAATGTCTCCAGACATCTTGATAATAC 1391  
Db |||||  
1329 TTTTGTGAGTGTGGTAGATACTCTGTCTACCAATGTCTATCCAGACATCTTGATAATAC 1388  
Qy |||||  
1392 CTGTTGGAATCTCCTATGATCGCAATTCGAAAGGTCACTAATATGTTGTAACAACTGGGCA 1451  
Db |||||  
1389 CTGTTGGAATCTCCTATGATCGCAATTCGAAAGGTCACTAATATGTTGTAACAACTGGGCA 1448  
Qy |||||

1452 AACCTAAGAAATGAGAGCCTGTGGAGTGTAGCAAGAGGTGTTATTAGAAATGTTACGAA 1511  
Db |||||  
1449 AACCTAAGAAATGAGAGCCTGTGGAGTGTAGCAAGAGGTGTTATTAGAAATGTTACGAA 1508  
Qy |||||  
1512 AAAAATATGTTGTGTCCGAGTGGATTTTGCACAGCCATTTTCTTAAAGGAATATTTAG 1571  
Db |||||  
1509 AAAAATATGTTGTGTCCGAGTGGATTTTGCACAGCCATTTTCTTAAAGGAATATTTAG 1568  
Qy |||||  
1572 AAAGCCAAAGTTCAGAAACGGTGTCTCTACTTTCTCCCTGGAGCAAGCGTGTCTTACCAG 1631  
Db |||||  
1569 AAAGCCAAAGTTCAGAAACGGTGTCTCTACTTTCTCCCTGGAGCAAGCGTGTGTACCAG 1628  
Qy |||||  
1632 CTATATCTCTTTCAAGACCAGTGTGCTGTGATGAAGGTAGAGACAGTCCATTAATG 1691  
Db |||||  
1629 CTATATCTCTTTCAAGACCAGTGTGCTGTGATGAAGGTAGAGACAGTCCATTAATG 1688  
Qy |||||  
1692 AGTCCAGAAATGCAACAGATGAATCCCTACGAAGAGGAGTGTGTTGCAAAATCTGGCTGAGC 1751  
Db |||||  
1689 AGTCCAGAAATGCAACAGATGAATCCCTACGAAGAGGAGTGTGTTGCAAAATCTGGCTGAGC 1748  
Qy |||||  
1752 ATATTCTATTTCACTGCTAGCAAGTCTGTGCCATTATGTCCACACATTTGGCTTGCC 1811  
Db |||||  
1749 ATATTCTATTTCACTGCTAGCAAGTCTGTGCCATTATGTCCACACATTTGGCTTGCC 1808  
Qy |||||  
1812 TGTCTCTCTACAGACACAGCGAGGGAATTTGATCTCTCCACATTTGGTTCGAAGACTTCTTTG 1871  
Db |||||  
1809 TGTCTCTCTACAGACACAGCGAGGGAATTTGATCTCTCCACATTTGGTTCGAAGACTTCTTTG 1868  
Qy |||||  
1872 TGATGAAGAGGAAGTCTTGGCTCGTGATTTTGAACCTGGGGTCTCAGGAAATTCAGAAG 1931  
Db |||||  
1869 TGATGAAGAGGAAGTCTTGGCTCGTGATTTTGAACCTGGGGTCTCAGGAAATTCAGAAG 1928  
Qy |||||  
1932 ATGTAGTAATGCATGCCATACAGCTGCTGGGAAATTTGTCTCAAAATTCACCCACACTAGCA 1991  
Db |||||  
1929 ATGTAGTAATGCATGCCATACAGCTGCTGGGAAATTTGTCTCAAAATTCACCCACACTAGCA 1988  
Qy |||||  
1992 GGAAATGATGAGTCTTTTATCACCCCGACACAACTGTCCCATCAGTCTTCGAATCAACT 2051  
Db |||||  
1989 GGAAATGATGAGTCTTTTATCACCCCGACACAACTGTCCCATCAGTCTTCGAATCAACT 2048  
Qy |||||  
2052 TCTACAGCAATGGGGTACTTTCATGTCTTTATCATGGAGGCCATCATAGCTTGAGCGCTTT 2111  
Db |||||  
2049 TCTACAGCAATGGGGTACTTTCATGTCTTTATCATGGAGGCCATCATAGCTTGAGCGCTTT 2108  
Qy |||||  
2112 ATGCGATCTGAACAGAGGGGACTGGGGGTCCCACTAGCACCCCACTTAACCTGATCA 2171  
Db |||||  
2109 ATGCGATCTGAACAGAGGGGACTGGGGGTCCCACTAGCACCCCACTTAACCTGATCA 2168  
Qy |||||  
2172 GCCAGAGCAGCTGGTGGGAAAGGGCCAGGCTGTGTACTCTTCTCCAATGAAGGCA 2231  
Db |||||  
2169 GCCAGAGCAGCTGGTGGGAAAGGGCCAGGCTGTGTACTCTTCTCCAATGAAGGCA 2228  
Qy |||||  
2232 CCATCTCACTGCTTGCACAGACATTTTACCAAGTCTGCCATGAACAGTAGGAAAGTTTA 2291  
Db |||||  
2229 CCATCTCACTGCTTGCACAGACATTTTACCAAGTCTGCCATGAACAGTAGGAAAGTTTA 2288  
Qy |||||  
2292 TCCAGTATGGCAATCTTACAGTGGCAGACGATGACAGGAGGAGATATCAGTCTTAGTC 2351  
Db |||||  
2289 TCCAGTATGGCAATCTTACAGTGGCAGACGATGACAGGAGGAGATATCAGTCTTAGTC 2348  
Qy |||||  
2352 TTGCTGAGCAGCTGGGCAAGAGAGCTTCCAGAACCTTTGTCTTGGAGAGTGAAG 2411  
Db |||||  
2349 TTGCTGAGCAGCTGGGCAAGAGAGCTTCCGAACTTTGTCTTGGAGAGTGAAG 2408  
Qy |||||  
2412 AAGATGAACAGTGTGCTTTGGGAGGAAACAGCGAGATTTGCTACCTGAAGGTGAGCCAAT 2471  
Db |||||  
2409 AAGATGAACAGTGTGCTTTGGGAGGAAACAGCGAGATTTGCTACCTGAAGGTGAGCCAAT 2468  
Qy |||||  
2472 CCAAGGAGCAGCAGAGTTTATCACTTCTTACAGAGACTCTCTGGGCTTTGCTGGAG 2531  
Db |||||  
2469 CCAAGGAGCAGCAGAGTTTATCACTTCTTACAGAGACTCTCTGGGCTTTGCTGGAG 2528  
Qy |||||  
2532 CTTACAGCTCTGCTGCCATCTTTTGTTCACAACTTTCAGTGGTCTCTTCCAGAACCTGAGT 2591



```
Db 2529 CCTACAGCTCTGCTGCCATCTTTGTTCAACATCTCAGTGGTCTGTTCCAGAACCTGAGT 2588
Qy 2592 ATCTGCAAAAGTTGACAAATACCTTAATAACAGAAACAGAAAGAAATGTTCCAGTATATG 2651
Db 2589 ATCTGCAAAAGTTGACAAATACCTTAATAACAGAAACAGAAAGAAATGTTCCAGTATATG 2648
Qy 2652 CTGAGAGTGCACATATTTGCTTTGTGGAAGATGCTGTGAAAATGTTTAAAGATATGGGG 2711
Db 2649 CTGAGAGTGCACATATTTGCTTTGTGGAAGATGCTGTGAAAATGTTTAAAGATATGGGG 2708
Qy 2712 TTTTCAAGGACCAACAAAGAGAGTCTGTTTGAAGTCTGTTTGAAGTCTGAGCAGCACTTTCTAC 2771
Db 2709 TTTTCAAGGACCAACAAAGAGAGTCTGTTTGAAGTCTGTTTGAAGTCTGAGCAGCACTTTCTAC 2768
Qy 2772 CTCAATGCAACCGACAAAACCTTCTAGAAATATATTTCTGAGTTTGTGTTGCTGTAGGTAA 2831
Db 2769 CTCAATGCAACCGACAAAACCTTCTAGAAATATATTTCTGAGTTTGTGTTGCTGTAGGTAA 2828
Qy 2832 CGTGTGGCACTGTGGCAAAATGAAGTCTAGATGAGTCTCTTGTAGTACCAAGCTTCT 2891
Db 2829 CGTGTGGCACTGTGGCAAAATGAAGTCTAGATGAGTCTCTTGTAGTACCAAGCTTCT 2888
Qy 2892 GGCTCAAGAGTTGAAGTGTCTTCCATGATCTGAGATCTGAGTCTGAGTCTGAGTCT 2951
Db 2889 GGCTCAAGAG-TTGAAGTGTCTTCCATGATCTGAGATCTGAGTCTGAGTCTGAGTCT 2946
Qy 2952 CCTGGAAGCAAGTGCCTTCTNCCCTCCATGATCTGAGATCTGAGTCTGAGTCT 3003
Db 2947 CCTGGAAGCAAGTGCCTTCTNCCCTCCATGATCTGAGATCTGAGTCTGAGTCT 2998
```

```
RESULT 3
HSM804404 4918 bp mRNA linear PRI 30-APR-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp451P0819 (from clone DKFZp451P0819);
DEFINITION complete cds.
ACCESSION AL833093
VERSION 1.0
KEYWORDS AL833093.2 GI:30268382
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 4918)
AUTHORS Bloembergen H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
Fobo G., Han M., Osanger A. and Wiemann S.
TITLE Direct Submission
JOURNAL Submitted (30-APR-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT On Apr 30, 2003 this sequence version replaced gi:21733684.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp451P0819) is available at the RZPD in Berlin.
Please contact the RZPD, Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1. 4918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="10q25.2"
/clone="DKFZp451P0819"
/tissue_type="cDNA-collection"
/clone_lib="451 (synonym: hlccl). Vector pSport1; host
DH108; sites NotI + SalI"
/dev_stage="adult"
1. 4918
gene
```

```
/gene="DKFZp451P0819"
200..2686
/gene="DKFZp451P0819"
/notes="glycerol-3-phosphate acyltransferase"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD89932.1"
/db_xref="GI:30268383"
/t_xref="GI:30268383"
RSATLWKESSMSRKPFFGVCYCTPQSDKFFNPSIPLGLRNVIYNETHTRH
GLWARRLSYVLFQERDVHKGMEATNTENVNSSRVOEALAEVAELNPDGSAQQS
KANKVAKKARILQENWATVSPAMIKLTGVLKLFNSFFNFQIHKGLQEMWKAAT
ETNLPGLFVHRSHIDYLLTLLTILFCHNIKAPIYASGNLNPFIPTLLHKLGCFI
RRLETPDRKDLVYRALLHGHIVELLROQFLEIPLGTRSRSGTSCARAGLLSV
VDTLTNTVPIIDILIIIPVGISYDRIIEGHVNGEOLGPKKNESLWSVARGVIRMLKVN
YGCVRVDFAPFSLKEVLESOKPSALLSLEQALIPALPSRPSDAADGGRDTSIN
ESRNATDESARRLLIANLAEHILFTASKSCAIMSTHIVACLLLYRHRQGLDSTLVED
FFVWKEVLRDLFDLFGNSDVMHAIQLGNCVTITHTSRNDEFTIDSTVPSV
FELNFSNGVLHVIMEAIACSLYAVLNKRGGLGPTSTPNLISQOLVRKAASLCY
LLNBSGTISLPCQTQVCHETVGFIRYGLITVAEHDDQEDISPISLAEQWKKLPE
PLWSRDEDEDDFGEORDCYLKVSQSEHOQIFLQRLGLPLLEAYSSAAIFVH
NFGSPVPEPYLOKLHKYLTIRTNVAVVAESATCYLVKNVKNFKDVGIVFKETKQK
RVSVLESLSTFLLPQCRKQLLEYILSPVL"
polyA_site 4893
BASE COUNT 1379 a 1005 c 1075 g 1459 t
ORIGIN
Query Match 91.8%; Score 2755.4; DB 9; Length 4918;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2781; Conservative 0; Mismatches 8; Indels 2; Gaps 2;
Qy 213 GGACAGAGCGGCTCCCTGTTGATGACATCTTGACCCGAAACTGATAGCTGAGTCT 272
Db 72 GGACAGAGCGGCTCCCTGTTGATGACATCTTGACCCGAAACTGATAGCTGAGTCT 131
Qy 273 GAAGTTTTATGTTATGAAACAGAGAACTTTTCATCCCGACACATGATTTGGGAATTACAC 332
Db 132 GAAGTTTTATGTTATGAAACAGAGAACTTTTCATCCCGACACATGATTTGGGAATTACAC 191
Qy 333 TTTGTGACATGGATGAATCTGACCTTGCTGTAACAATAGATGTTTCTTATCTGCCAC 392
Db 192 TTTGTGACATGGATGAATCTGACCTTGCTGTAACAATAGATGTTTCTTATCTGCCAC 251
Qy 393 ATTCAATCAGAAATACAGTGTGTTGTCGATGTAAGACACACAAAGTAGGAATGGGTGAGTGTG 452
Db 252 ATTCAATCAGAAATACAGTGTGTTGTCGATGTAAGACACACAAAGTAGGAATGGGTGAGTGTG 311
Qy 453 GCTTTAGACCCACCGCTCTTCAGATCTGCAACTTTTAAATGAAAGAAAGCCTAATGAGTC 512
Db 312 GCTTTAGACCCACCATCTTCAGATCTGCAACTTTTAAATGAAAGAAAGCCTAATGAGTC 371
Qy 513 GGAAGAGCCATTTGTTGGAAGATGTTGTTATCTCTGACCTCCCGACAGCTGGGACAAAT 572
Db 372 GGAAGAGCCATTTGTTGGAAGATGTTGTTATCTCTGACCTCCCGACAGCTGGGACAAAT 431
Qy 573 TTTTCAACCCACGATATCCCTCTTTGGTGTGCGGAATGTTATTTATCAATCAAACTC 632
Db 432 TTTTCAACCCACGATATCCCTCTTTGGTGTGCGGAATGTTATTTATCAATCAAACTC 491
Qy 633 ACACAAGACACCCGGATGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGAGC 692
Db 492 ACACAAGACACCCGGATGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGAGC 551
Qy 693 GAGATGTCATAAGGGCATGTTTGGCCACCAATGTGACTGAAAAATGTCTGAACAGCAGTA 752
Db 552 GAGATGTCATAAGGGCATGTTTGGCCACCAATGTGACTGAAAAATGTCTGAACAGCAGTA 611
Qy 753 GAGTACAGAGGCAATGTCAGAGTGGCTGCTGAATTAACCCCTGATGGTGTCTGCCAGC 812
Db 612 GAGTACAGAGGCAATGTCAGAGTGGCTGCTGAATTAACCCCTGATGGTGTCTGCCAGC 671
Qy 813 AGCAATCAAAAGCGGTTTAAACAAAGTGAAGAAAGCTAAAGAGTATTTCTTCAAGAAATGG 872
```

Db	672		AGCAATCAAAAGCGTTTAAACAAAGTGAAGAAAGCTAAAGGAATCTTCAAGAAATGG	731
Qy	873		TTGCCATGCTCACCAGCAATGATCAGACTGACTGGTGGTGGCTTAAACCTGTCA	932
Db	732		TTGCCATGCTCACCAGCAATGATCAGACTGACTGGTGGTGGCTTAAACCTGTCA	791
Qy	933		ACAGCTCTTTTGGAAACATTCAAATTCACAAAGGTCAATTTGAGATGGTTAAAGCTGCAA	992
Db	792		ACAGCTCTTTTGGAACTTTCAATTCACAAAGGTCAATTTGAGATGGTTAAAGCTGCAA	851
Qy	993		CTGAGACGAATTTGCCGCTCTCTGTTCTACCAAGTTCATAGATCCCATATTTGACTATCTGC	1052
Db	852		CTGAGACGAATTTGCCGCTCTCTGTTCTACCAAGTTCATAGATCCCATATTTGACTATCTGC	911
Qy	1053		TGCTCATTCTCTCTCTGCTGCCATAACATCAAGACCATATGCTTCAGGCAATA	1112
Db	912		TGCTCATTCTCTCTCTGCTGCCATAACATCAAGACCATATGCTTCAGGCAATA	971
Qy	1113		ATCTCAACATCCCAATCTTCAGTACCTTTGATCCATAGCTTTGGGGCTCTTCTCATAGCAC	1172
Db	972		ATCTCAACATCCCAATCTTCAGTACCTTTGATCCATAGCTTTGGGGCTCTTCTCATAGCAC	1031
Qy	1173		GAAGGCTCGATGAACACCCAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATG	1232
Db	1032		GAAGGCTCGATGAACACCCAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATG	1091
Qy	1233		GGCATATAGTTGAATTTCTGACAGCAGCAATTTCTGGAGATCTTCTGGAAAGGCACAC	1292
Db	1092		GGCATATAGTTGAATTTCTGACAGCAGCAATTTCTGGAGATCTTCTGGAAAGGCACAC	1151
Qy	1293		GTCTAGAGTGGAAACCTTTGTGCTCGGCGAGGACTTTTGTGACATTTGTGCTGATGA	1352
Db	1152		GTCTAGAGTGGAAACCTTTGTGCTCGGCGAGGACTTTTGTGACATTTGTGCTGATGA	1211
Qy	1353		CTCTGTCTACCAATGTCATCCAGACATCTTGATAATACCTGTTGGAATCTCCTATGATC	1412
Db	1212		CTCTGTCTACCAATGTCATCCAGACATCTTGATAATACCTGTTGGAATCTCCTATGATC	1271
Qy	1413		GCATTATCGAAGTCACTAAATGGTGAACAACTGGGCAACCTTAGAAGAAATGAGAGCC	1472
Db	1272		GCATTATCGAAGTCACTAAATGGTGAACAACTGGGCAACCTTAGAAGAAATGAGAGCC	1331
Qy	1473		TGTGGAGTGAAGAGGTTTATTAAGATGTTTACGAAACAACTATGTTGTCGCGAG	1532
Db	1332		TGTGGAGTGAAGAGGTTTATTAAGATGTTTACGAAACAACTATGTTGTCGCGAG	1391
Qy	1533		TGGATTTTGACAGCCATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGG	1592
Db	1392		TGGATTTTGACAGCCATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGG	1451
Qy	1593		TGTCGTCTATCTTCTCGAGCAAGCGTTGTTTACAGCTATATCTTCTTCAAGACCCA	1652
Db	1452		TGTCGTCTATCTTCTCGAGCAAGCGTTGTTTACAGCTATATCTTCTTCAAGACCCA	1511
Qy	1653		GTGATGCTGTGATGAGAGTGAAGACAGCTCCATTAATAGTCCAGAAATGCAACAGATG	1712
Db	1512		GTGATGCTGTGATGAGAGTGAAGACAGCTCCATTAATAGTCCAGAAATGCAACAGATG	1571
Qy	1713		AATCCCTACGAAGAGGTTGATTTGCAATCTGGCTGAGCATATTTCTATTCACCTGCTAGCA	1772
Db	1572		AATCCCTACGAAGAGGTTGATTTGCAATCTGGCTGAGCATATTTCTATTCACCTGCTAGCA	1631
Qy	1773		AGTCCTGTGCCATTAATGTCACACACATTTGGCTGGCTGCTCTCTACAGACAGGC	1832
Db	1632		AGTCCTGTGCCATTAATGTCACACACATTTGGCTGGCTGCTCTCTACAGACAGGC	1691
Qy	1833		AGGGAATTTGATCTCTCCACATTTGTCGAAGACTTTCTTTGATGAAGAGAGAGTCTCTGG	1892
Db	1692		AGGGAATTTGATCTCTCCACATTTGTCGAAGACTTTCTTTGATGAAGAGAGAGTCTCTGG	1751
Qy	1893		CTCGTGAATTTGACCTGGGGTCTCTCAGGAAATTCAGAAAGATAGTAATGCATGCCATAC	1952

Db	1752		CTCTGATTTTGCACCTGGGGTTCTCAGGAAATTCAGAAAGATAGTAATGCATGCCATAC	1811
Qy	1953		AGCTGCTGGGAAATTTGTGTCAAAATCACCCACACTAGCAGGAAACGATGAGTTTTTATCA	2012
Db	1812		AGCTGCTGGGAAATTTGTGTCAAAATCACCCACACTAGCAGGAAACGATGAGTTTTTATCA	1871
Qy	2013		CCCCAGCACAACCTGTCCCATCAGTCTTCGAACTCAAACTTTACAGCAATGGGTACTTTC	2072
Db	1872		CCCCAGCACAACCTGTCCCATCAGTCTTCGAACTCAAACTTTACAGCAATGGGTACTTTC	1931
Qy	2073		ATGCTCTTATCATGAGAGCCCATCATAGCTTTCAGAGCTTTTATGAGTTCTGAAACAGAGGG	2132
Db	1932		ATGCTCTTATCATGAGAGCCCATCATAGCTTTCAGAGCTTTTATGAGTTCTGAAACAGAGGG	1991
Qy	2133		GACTGGGGTCCCACTAGCACCCCACTAACTGATCAGCCAGCAGCAGCTGTGCGGA	2192
Db	1992		GACTGGGGTCCCACTAGCACCCCACTAACTGATCAGCCAGCAGCAGCTGTGCGGA	2051
Qy	2193		AGGCGGCCAGGCTGTGCTACCTTCTCTCAATGAAGGCAACCATCTCAGCTTGCCTTGCAGA	2252
Db	2052		AGGCGGCCAGGCTGTGCTACCTTCTCTCAATGAAGGCAACCATCTCAGCTTGCCTTGCAGA	2111
Qy	2253		CATTTTACCAAGTGTGCCATGAAGAACATGAGGAAAGTTTATCCAGTATGCGATTTTACAG	2312
Db	2112		CATTTTACCAAGTGTGCCATGAAGAACATGAGGAAAGTTTATCCGGTATGCGATTTTACAG	2171
Qy	2313		TGCGAGCAGCAGTACAGCAGGAGATATCAGTCTTAGTCTCTGCTGAGCAGCAGTGGGACA	2372
Db	2172		TGCGAGCAGCAGTACAGCAGGAGATATCAGTCTTAGTCTCTGCTGAGCAGCAGTGGGACA	2231
Qy	2373		AGAAGCTTCCAGAACCTTTGTCTTGAGAAAGTGATGAAGAAATGGAAGACAGTGAATTTG	2432
Db	2232		AGAAGCTTCCAGAACCTTTGTCTTGAGAAAGTGATGAAGAAATGGAAGACAGTGAATTTG	2291
Qy	2433		GGGAGAAACAGCAGAGATGCTACCTGAAGGTGAGCAATCCAAAGGAGCACAGCAGTTTA	2492
Db	2292		GGGAGAAACAGCAGAGATGCTACCTGAAGGTGAGCAATCCAAAGGAGCACAGCAGTTTA	2351
Qy	2493		TCACCTTCTTACAGAGCTCTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCT	2552
Db	2352		TCACCTTCTTACAGAGCTCTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCT	2411
Qy	2553		TTGTTTCAACAACTTCAAGTGTCTGTTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAAT	2612
Db	2412		TTGTTTCAACAACTTCAAGTGTCTGTTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAAT	2471
Qy	2613		ACCTAATAACAGAAACAGAAAGAAATGTTGCAAGTATATGCTGAGAGTGCCATATTTGC	2672
Db	2472		ACCTAATAACAGAAACAGAAAGAAATGTTGCAAGTATATGCTGAGAGTGCCATATTTGC	2531
Qy	2673		TTGTGAGAAATGCTGTGAAATGTTTAAAGTATTTGGGGTTTTCAAGGAGACCAACAA	2732
Db	2532		TTGTGAGAAATGCTGTGAAATGTTTAAAGTATTTGGGGTTTTCAAGGAGACCAACAA	2591
Qy	2733		AGAGAGTGTCTGTTTGTAGAACTGAGCAGCCTTTTCTACCTCAATGCAACCCGACCAAAAC	2792
Db	2592		AGAGAGTGTCTGTTTGTAGAACTGAGCAGCCTTTTCTACCTCAATGCAACCCGACCAAAAC	2651
Qy	2793		TTCTAGAAATATATCTGAGTTTTTGTGGTCTGTAGGTAAACGTTGGCACTGTGTGGCAAT	2852
Db	2652		TTCTAGAAATATATCTGAGTTTTTGTGGTCTGTAGGTAAACGTTGGCACTGTGTGGCAAT	2711
Qy	2853		GAAGTTCATGAGATGAGTTTCTTGTAGGTACAGCTTCTGCTCAGAGTTTGAAGGTGC	2912
Db	2712		GAAGTTCATGAGATGAGTTTCTTGTAGGTACAGCTTCTGCTCAGAGTTTGAAGGTGC	2770
Qy	2913		CTTTCAGAGGGGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2972
Db	2771		CATGCA - GGGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2829
Qy	2973		NCCTCATGATCTGAGATCTTCCACAGCTTT	3003
Db	2830		CCCTCATGATCTGTGATCTTCCACAGCTCT	2860

RESULT 4  
BC030783  
LOCUS  
DEFINITION Homo sapiens, KIAA1560 protein, clone MGC:26846 IMAGE:4821342,  
3490 bp mRNA linear PRI 03-JUN-2002  
mRNA, complete cds.  
ACCESSION BC030783  
VERSION BC030783.1 GI:21315057  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Strausberg, R.  
Direct Submission  
Submitted (31-MAY-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
Contact: [amandawebster@systemsbiology.org](mailto:amandawebster@systemsbiology.org)  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha  
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAP Plate: 33 Row: b Column: 19  
This clone was selected for full length sequencing because it  
passed the following selection criteria: GenomesCan gene  
prediction, Similarity but not identity to protein.  
FEATURES  
Location/Qualifiers  
1..3490  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/db\_xref="LocusID:57678"  
/clone="MGC:26846 IMAGE:4821342"  
/cissue\_type="Brain, hippocampus"  
/clone\_lib="NIH MGC\_95"  
/lab\_host="DH10B"  
/note="Vector: pBluescript"  
179..2665  
/codon\_start=1  
/product="KIAA1560 protein"  
/protein\_id="AAH30783.1"  
/db\_xref="GI:21315058"  
/translation="MDEALTLGTDVSLPHSSEYSGVCKHTSEWGEGRPTIP  
RSATLKWKESLSMRKPEVGCYCTPQSDWKFTNTSIPSLGLRNVIYNEHTRH  
GWLARRLSYLFIGERDVHKNFPAINTVENSRSQEAIAEVAELNPDGSAQQS  
KAVNKVKAKRILOEMVATVSPAMIRLTGWLLKLFNSFFWNIQIHKGQLEVMKAA  
ETNPLFLPVHRSHIDYLLTFLFCHNIKAPYIASGNLNIPISFTLHKLGFFI  
RRLEDEYDGRKDVLYRALLHGHIVELLRQOQFLEIFLEGRSTRSGKTSARAGLLV  
VVDLTSTNVI PDILIPVGISYDRI IEHNGEOLGPKKNESLWSVARGVRLRN  
YGCVRVPAQPSFLSKVLESQKPSVALLSLEQALLPALIPSPSDAEDGRSTIN  
ESRNATESLRRLLIANLAELHIFTSKCAIMSTHIVACLLLYRRHQGLDLSLVED  
FVWKVEELDFDGLSGNSDEDVVMHAIQLLGNCTVITHTSRNDEFFITPSTVPSV  
FELNPLFNGVLTQVIMEAIIACSLYAVLNKRGLGPTSTPENLISQBLVRKAASLCY  
LLSNEGITSLPCQTFYOVCHETGVGFIOYGLITVAEHDDQEDISPLAEOQWKKLPE  
PLSKRSDEEDSDYGEORDCYLVKSQSEKHQOQFIFLQRLGLGLEAYSSAIFVH  
NFGSPVPEYVLOKLKXLYLTRERNVAVYAESATYCLVRNVAQKFDIGVFKETKQ  
RVSVLELSTFLPQCNRQKQLLEYILISFVLL"  
1002 a 734 c 790 g 964 t  
BASE COUNT  
ORIGIN

Query Match 91.6%; Score 2752.2; DB 9; Length 3490;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 2779; Conservative 0; Mismatches 10; Indels 2; Gaps 2;  
QY 213 GGCAGCAGCGCTCCCTCTGTTGTATGACATCTTCGACCCGAACTGATAGCTAGTCT 272  
DB 51 GGCAGCAGCGCTCCCTCTGTTGTATGACATCTTCGACCCGAACTGATAGCTAGTCT 110  
QY 273 GAAGTTTTATGTTATGAAACAGAGAACTTTTCATCCCGACGACATGATTGGGAATTACAC 332  
DB 111 GAAGTTTTATGTTATGAAACAGAGAACTTTTCATCCCGACGACATGATTGGGAATTACAC 170  
QY 333 TTTGTGACATGGATGAATCTGCACTGACCCCTTGGTCAACAATAGATGTTTCTTATCTGCCAC 392  
DB 171 TTTGTGACATGGATGAATCTGCACTGACCCCTTGGTCAACAATAGATGTTTCTTATCTGCCAC 230  
QY 393 ATTCAATCAGATACAGTGTGGTGCATGTAAAGACACACAGTGGAGAAATGGGGTGAAGTG 452  
DB 231 ATTCAATCAGATACAGTGTGGTGCATGTAAAGACACACAGTGGAGAAATGGGGTGAAGTG 290  
QY 453 GCTTTAGACCCACCGTCTTCAGATCTCAACTTTTAAATGGAAGAAAGCCCTAATAGATC 512  
DB 291 GCTTTAGACCCACCGTCTTCAGATCTCAACTTTTAAATGGAAGAAAGCCCTAATAGATC 350  
QY 513 GGAAGAGCCCATTTGTTGGAAGATGTTGTTACTCTCTGCACCTCCCGACAGCTGGGACAAAT 572  
DB 351 GGAAGAGCCCATTTGTTGGAAGATGTTGTTACTCTCTGCACCTCCCGACAGCTGGGACAAAT 410  
QY 573 TTTTCAACCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAACTC 632  
DB 411 TTTTCAACCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAACTC 470  
QY 633 ACACAGACACCGCGATGGCTTGCAGAGCCCTTCTTACGTTCTTTTATTTCAAGAGC 692  
DB 471 ACACAGACACCGCGATGGCTTGCAGAGCCCTTCTTACGTTCTTTTATTTCAAGAGC 530  
QY 693 GAGATGTGCATAGGGCATGTTTGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTA 752  
DB 531 GAGATGTGCATAGGGCATGTTTGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTA 590  
QY 753 GAGTACAAGAGGCAATTCGAGAAGTGGCTGCTGAATTTAAACCCCTGATGGTTCCTCCACG 812  
DB 591 GAGTACAAGAGGCAATTCGAGAAGTGGCTGCTGAATTTAAACCCCTGATGGTTCCTCCACG 650  
QY 813 AGCAATCAAAAGCGTTTAAACAAAGTGAAGAAAGCTTAAAGCAATCTTCAAGAAATGG 872  
DB 651 AGCAATCAAAAGCGTTTAAACAAAGTGAAGAAAGCTTAAAGCAATCTTCAAGAAATGG 710  
QY 873 TTGCCACTGTCTCACCGGCAATGATCAGATGCTGCTGGGTGGGTGCTGCTAAACCTGTTCA 932  
DB 711 TTGCCACTGTCTCACCGGCAATGATCAGATGCTGCTGGGTGGGTGCTGCTAAACCTGTTCA 770  
QY 933 ACAGCTTTCTTTTGGAAATTCAAATTCACAAAGTCAACTTTGAGATGGTGTAAAGCTGCAA 992  
DB 771 ACAGCTTTCTTTTGGAAATTCAAATTCACAAAGTCAACTTTGAGATGGTGTAAAGCTGCAA 830  
QY 993 CTGAGACGAATTTGCCCGCTTCTGTTCTACAGTTTCATAGATCCCATATTGACTATCTGC 1052  
DB 831 CTGAGACGAATTTGCCCGCTTCTGTTCTTACAGTTTCATAGATCCCATATTGACTATCTGC 890  
QY 1053 TGCTCACTTTTCATTTCTCTTCTGCCATAACCAAGACCAATACATTTGCTTCAGGCAATA 1112  
DB 891 TGCTCACTTTTCATTTCTCTTCTGCCATAACCAAGACCAATACATTTGCTTCAGGCAATA 950  
QY 1113 ATCTCAACATCCCAATCTTCAGTACCTTGTATCCATAAGCTTTGGGGGCTTCTTCATAGCAG 1172  
DB 951 ATCTCAACATCCCAATCTTCAGTACCTTGTATCCATAAGCTTTGGGGGCTTCTTCATAGCAG 1010  
QY 1173 GAAGGCTCGATGAACACACAGATGGAGGAAAGATGTTCTCTATAGAGCTTTGCTCCATG 1232  
DB 1011 GAAGGCTCGATGAACACACAGATGGAGGAAAGATGTTCTCTATAGAGCTTTGCTCCATG 1070

Qy	1233	GGCATATAGTTGAAATTA	CTTCGACAGAGCAAA	TTCTTTGGAGATCTT	CCTCGAAGGCACAC	1292	
Db	1071	GGCATATAGTTGAAATTA	CTTCGACAGAGCAAA	TTCTTTGGAGATCTT	CCTCGAAGGCACAC	1130	
Qy	1293	GTTCCTAGGAGTGGAAAAC	CTCTTGCTCGGCGAGCAG	CTTTTGTCAAGTTGT	GTGAGTGTGGTAGATA	1352	
Db	1131	GTTCCTAGGAGTGGAAAAC	CTCTTGCTCGGCGAGGAC	CTTTTGTCAAGTTGT	GTGAGTGTGGTAGATA	1190	
Qy	1353	CTCTGTCTACCAATGTG	CATCCCAGACATCTTG	GATAAATACCTGTTGG	GAATCTCTTATGATC	1412	
Db	1191	CTCTGTCTACCAATGTG	CATCCCAGACATCTTG	ATAAATACCTGTTGG	GAATCTCTTATGATC	1250	
Qy	1413	GCATTTATCGAAGGTCA	CTACAAATGGTGAAACA	ACTGGGCAAACTTA	GAGAAATGAGAGCC	1472	
Db	1251	GCATTTATCGAAGGTCA	CTACAAATGGTGAAACA	ACTGGGCAAACTTA	GAGAAATGAGAGCC	1310	
Qy	1473	TGTGGAGTGTAGCAAGAG	GTGTTATTAGATGTTT	ACGAAAATACTATGG	TTGTGTCCGAG	1532	
Db	1311	TGTGGAGTGTAGCAAGAG	GTGTTATTAGATGTTT	ACGAAAATACTATGG	TTGTGTGTCCGAG	1370	
Qy	1533	TGGATTTTGCACAGCCAT	TTTTCTTTAAAGGAATA	TTTAAAGGCCAAAGT	CAGAAAACCGG	1592	
Db	1371	TGGATTTTGCACAGCCAT	TTTTCTTTAAAGGAATA	TTTAAAGGCCAAAGT	CAGAAAACCGG	1430	
Qy	1593	TGTCGTCTCTACTTTT	CCCTCGAGCAAGCTT	GTGTACCACTATACT	TTCTTCAAGACCCA	1652	
Db	1431	TGTCGTCTCTACTTTT	CCCTCGAGCAAGCTT	GTGTACCACTATACT	TTCTTCAAGACCCA	1490	
Qy	1653	GTGATGCTGCTGATGA	AGGTAGACACAGTCC	ATTAATGAGTCCAGA	AAATGCAACAGATG	1712	
Db	1491	GTGATGCTGCTGATGA	AGGTAGACACAGTCC	ATTAATGAGTCCAGA	AAATGCAACAGATG	1550	
Qy	1713	AATCCCTACGAAGGAGT	GATTGTCAAATCTTGG	CTGAGCATATTCTAT	TCTCACTTGCCTAGCA	1772	
Db	1551	AATCCCTACGAAGGAGT	GATTGTCAAATCTTGG	CTGAGCATATTCTAT	TCTCACTTGCCTAGCA	1610	
Qy	1773	AGTCCTGTGCCATTAT	GTGCCACACACATTT	GTGGCTTGCCCTG	CTCTACAGACACAGGC	1832	
Db	1611	AGTCCTGTGCCATTAT	GTGCCACACACATTT	GTGGCTTGCCCTG	CTCTACAGACACAGGC	1670	
Qy	1833	AGGGAATTGATCTCT	CCACATTTGTCGAGA	CACTTTCTTTGTGTA	GAAAGAGAGTCTCTG	1892	
Db	1671	AGGGAATTGATCTCT	CCACATTTGTCGAGA	CACTTTCTTTGTGTA	GAAAGAGAGTCTCTG	1730	
Qy	1893	CTCGTGATTTTGACCT	CGGGTTCTCAGSAA	TTCCAGAAAGTGTAG	TAAATGCATGCCATAC	1952	
Db	1731	CTCGTGATTTTGACCT	CGGGTTCTCAGSAA	TTCCAGAAAGTGTAG	TAAATGCATGCCATAC	1790	
Qy	1953	AGCTGTGGGAAATTT	GTGTACAAATCACCA	CACTAGCAGGAAAC	GATGAGTTTTTTATCA	2012	
Db	1791	AGCTGTGGGAAATTT	GTGTACAAATCACCA	CACTAGCAGGAAAC	GATGAGTTTTTTATCA	1850	
Qy	2013	CCCCAGACAAATGTC	ATGCCATCTGAACTT	CGAACTTCTACAG	CAATGGGGTACTTC	2072	
Db	1851	CCCCAGACAAATGTC	ATGCCATCTGAACTT	CGAACTTCTACAG	CAATGGGGTACTTC	1910	
Qy	2073	ATGTCTTTTATCATG	GAGGCCATCATAGCT	TTCGACCTTTATG	CAGTCTCTGAAAC	AAAGAGG	2132
Db	1911	ATGTCTTTTATCATG	GAGGCCATCATAGCT	TTCGACCTTTATG	CAGTCTCTGAAAC	AAAGAGG	1970
Qy	2133	GACTGGGGGGTCCCA	CTAGCACCCCACT	AACTTGATCAGCC	AGGAGCAGCTGGTGCGGA	2192	
Db	1971	GACTGGGGGGTCCCA	CTAGCACCCCACT	AACTTGATCAGCC	AGGAGCAGCTGGTGCGGA	2030	
Qy	2193	AGGCGGCCAGCTGTG	TGCTACTCTCCAA	TGAAGGCA	CAATCTCACTGCTTGC	CAAG	2252
Db	2031	AGGCGGCCAGCTGTG	TGCTACTCTCCAA	TGAAGGCA	CAATCTCACTGCTTGC	CAAG	2090
Qy	2253	CATTTTACCAAGTCTG	CCATGAACAGTGG	AAAGTTTATCCAG	TATGCAATCTTCTAC	G	2312
Db	2091	CATTTTACCAAGTCTG	CCATGAACAGTGG	AAAGTTTATCCAG	TATGCAATCTTCTAC	G	2150
Qy	2313	TGGCAGAGCAATGAC	CAGGAAGATATCAG	TCTTCTGTGAC	GACAGCAGTGGGACA	2372	

[illegible]

RESULT 5  
HSM804372  
LOCUS  
DEFINITION  
HSM804372 6387 bp mRNA linear PRI 30-APR-2003  
Homo sapiens mRNA; cDNA DKFP451B1115 (from clone DKFP451B1115);  
complete cds.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 6387)  
Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Weill,B., Amid,C.,  
Fobo G. Han, M., Osanger,A. and Wiemann,S.  
AUTHORS  
Direct Submission  
TITLE  
Submitted (30-APR-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANY  
JOURNAL  
On Apr 30, 2003 this sequence version replaced gi:21733651.  
COMMENT  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by GGF (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
This clone (DKFP451B1115) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES		Location/Qualifiers	
source	1..6387	/organism="Homo sapiens"	
		/mol_type="mRNA"	
	/db_xref="taxon:9606"		
	/map="10q25.2"		
	/clone="DKFZp451B1115"		
	/tissue_type="cDNA-collection"		
	/clone_lib="451 (synonym: hlccl). Vector pSport1; host DH10B; sites NotI + SalI"		
	/dev_stage="adult"		
gene	1..6387	/gene="DKFZp451B1115"	
	197..2683		
CDS	/gene="DKFZp451B1115"		
	/note="glycerol-3-phosphate acyltransferase"		
	/codon_start=1		
	/product="hypothetical protein"		
	/protein_id="CAD89967.1"		
	/db_xref="GI:30268347"		
	/translation="WDSEALTGLTIDVSLPHSSEYSVGRCKHTSEWGECCGFRPTVP RSATLWKESLMRSRRPVGRCYSCCTPQSDWDFNPSIPSLGRNVIIYINETHRHR GMLARRLSYLFIOERDVHKGMPATNTENVLSNRVQEAIAEVAEALNPDGSAQQQS KAVNKVKKAKRIIQEMVATVSPAMIRLTGMWLKLFNSPFNNQIHKQLEMYKAAT ETNLLPLFVHRSHIDYLLTFLFCHNIKAPVIASNNNLIPIFSTLHLKGGFT RRLEDETGDGRKDVLYRALLGHVLELLRQQQFLIEIFLETRSRSKTSRCARGLLSV VYDTLTPDIPDLIPGISYRIEGHYNGEQKPKNESLWSVARGVIRWLRKN YGCVRVDFAPQIFLSEYQSQKPSVALLSLEQALLPAILPSRPSDAADRDTNIN ESRNATDESRLRIANLAHEILFTASKSCAIMSTHIVACILLYRHRQGLDLSTLVED PFVKEEVLARDFLDGFSGNEDVVMHAIQLGNCVTITHTSRNDEPFIPTSTVPSV PELNFYSNGLHVITMEAI IACSLYAVLNKEGLGPTSPENLISQQLVRKAASLCY LLSNEATISLCQFQYQVCHETVKFQYQGLITVAEHDDQEDISFLAEQWKKGP LLSWRDESDSDSDFGEQRDCYILKVSQSEKHQFIFLORLLGLLLEAYSSAAIFVH NFSQVEPEYELQKLHKYLIITERNVAVAESATYCLVKNVAMFKDIGVFKETQK RVSVLELSSTFLPQCNROKLELYLSFVVL"		
	6346..6351		
polyA_signal	/gene="DKFZp451B1115"		
	6364		
polyA_site	/gene="DKFZp451B1115"		
	1865 a 1239 c 1381 g 1902 t		
BASE COUNT			
ORIGIN			
Query Match		91.6%; Score 2752.2; DB 9; Length 6387;	
Best Local Similarity		99.6%; Pred. No. 0;	
Matches 2779; Conservative		0; Mismatches 10; Indels 2; Gaps 2;	
Qy	213	GGCAGCAGCGCTCCCGCTGTGTATGGACATCTGACCCGAACTGATAGCTGAGTCCT	272
Db	69	GGCAGCAGCGCTCCCGCTGTGTATGGACATCTGACCCGAACTGATAGCTGAGTCCT	128
Qy	273	GAACTTTTATGTTATGAACAGAGAACTTTTCATCCGACGACATGATTGGGAATTACAC	332
Db	129	GAACTTTTATGTTATGAACAGAGAACTTTTCATCCGACGACATGATTGGGAATTACAC	188
Qy	333	TTTGTGACATGATGAATCTGACGACCTTGGTACATGATGTTCTTATCTGCCAC	392
Db	189	TTTGTGACATGATGAATCTGACGACCTTGGTACATGATGTTCTTATCTGCCAC	248
Qy	393	ATTTCATCAGAAATACAGTGTGGTCCATGTAAGCACAAGTGAGGAATGGGGTGGTGTG	452
Db	249	ATTTCATCAGAAATACAGTGTGGTCCATGTAAGCACAAGTGAGGAATGGGGTGGTGTG	308
Qy	453	GCTTTAGACCCCGCTCTTCAGATCTGCAACTTTAAATGGAAAGAAAGCCTAATGAGTC	512
Db	309	GCTTTAGACCCCGCTCTTCAGATCTGCAACTTTAAATGGAAAGAAAGCCTAATGAGTC	368
Qy	513	GGAAAGGCCCAATTTGTTGAGAGATGTTGTTACTCTGCACTCCCGACAGCTGGGACAAAT	572
Db	369	GGAAAGGCCCAATTTGTTGAGAGATGTTGTTACTCTGCACTCCCGACAGCTGGGACAAAT	428

573 TTTTCAACCCCGAGTATCCCGTCTTTTGGGTTTGGCGAATGTTATTATATCAATGAAATC 632  
429 TTTTCAACCCCGAGTATCCCGTCTTTTGGGTTTGGCGAATGTTATTATATCAATGAAATC 488  
633 ACACAAGACACCCCGGATGGCTTCCCAACCAATGTGACTGAAAAATGTGCTGAACAGCAGTA 692  
489 ACACAAGACACCCCGGATGGCTTCCCAACCAATGTGACTGAAAAATGTGCTGAACAGCAGTA 548  
693 GAGATGTGCAATAAGGGCATGTTTCCCAACCAATGTGACTGAAAAATGTGCTGAACAGCAGTA 752  
549 GAGATGTGCAATAAGGGCATGTTTCCCAACCAATGTGACTGAAAAATGTGCTGAACAGCAGTA 608  
753 GAGTACAAGAGGCAATTTGCAGAAAGTGGCTGCTGAAATTAACCCCTGATGGTTCTGCCACAGC 812  
609 GAGTACAAGAGGCAATTTGCAGAAAGTGGCTGCTGAAATTAACCCCTGATGGTTCTGCCACAGC 668  
813 AGCAATCAAAAGCCGTTAAACAAAGTGAAAAAGAAAGCTAAAGGATCTTCAAGAAATGG 872  
669 AGCAATCAAAAGCCGTTAAACAAAGTGAAAAAGAAAGCTAAAGGATCTTCAAGAAATGG 728  
873 TTGCACCTGTCTACCCGGCAATGATCAGACTGACTGCTGGTGGTGGTCTGCTAAACCTGTTCA 932  
729 TTGCACCTGTCTACCCGGCAATGATCAGACTGACTGCTGGTGGTGGTCTGCTAAACCTGTTCA 788  
933 ACAGCTTCTTTTGGAAACATTTCAAAATTCACAAAGGTCAACTTTGAGATGGTTAAAGCTGCAA 992  
789 ACAGCTTCTTTTGGAAACATTTCAAAATTCACAAAGGTCAACTTTGAGATGGTTAAAGCTGCAA 848  
993 CTGAGAGCAATTTGCCCGTCTCTGTTCTTACAGTTTCATAGATCCCATATGACTATCTGC 1052  
849 CTGAGAGCAATTTGCCCGTCTCTGTTCTTACAGTTTCATAGATCCCATATGACTATCTGC 908  
1053 TGCTCAGTTTCATCTCTCTGCCATAACATCAAGCACCATATACATTCGTTTCAGGCAATA 1112  
909 TGCTCAGTTTCATCTCTCTGCCATAACATCAAGCACCATATACATTCGTTTCAGGCAATA 968  
1113 ATCTCAACATCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGGCTTCTTCATACGAC 1172  
969 ATCTCAACATCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGGCTTCTTCATACGAC 1028  
1173 GAAGCTCGATGAAACACACAGATGGACGGAAGATGTTCTATAGAGCTTTGTCTCCATG 1232  
1029 GAAGCTCGATGAAACACACAGATGGACGGAAGATGTTCTATAGAGCTTTGTCTCCATG 1088  
1233 GGCATATAGTTGAATTAATCTTCGACGACGACCAATTTCTTGGAGATCTTCTGGAAGGCACAC 1292  
1089 GGCATATAGTTGAATTAATCTTCGACGACGACCAATTTCTTGGAGATCTTCTGGAAGGCACAC 1148  
1293 GTTCTAGGAGTGGAAAAACCTTGTGCTCGGGCAGGACATTTTGTCTAGTTGTGGTAGATA 1352  
1149 GTTCTAGGAGTGGAAAAACCTTGTGCTCGGGCAGGACATTTTGTCTAGTTGTGGTAGATA 1208  
1353 CTCTGTCTCAATGTCTATCCAGACATCTTTGATAATACCTTGTGGAAATCTCTTATGATC 1412  
1209 CTCTGTCTCAATGTCTATCCAGACATCTTTGATAATACCTTGTGGAAATCTCTTATGATC 1268  
1413 GCATTATCGAAGTCACTACAAATGGTGCAACAACTGGGGCAAACTTAAGAAGATGAGAGCC 1472  
1269 GCATTATCGAAGTCACTACAAATGGTGCAACAACTGGGGCAAACTTAAGAAGATGAGAGCC 1328  
1473 TGTGAGTGTAGCAAGAGGTGTTATTAGAAATGTTTACGAAAAAATATGTTGTGTCTCCGAG 1532  
1329 TGTGAGTGTAGCAAGAGGTGTTATTAGAAATGTTTACGAAAAAATATGTTGTGTCTCCGAG 1388  
1533 TGGATTTTGCACGCCAATTTTCTTAAAGGAATAATTTAGAAAGCCAAAGTCAGAAACCGG 1592  
1389 TGGATTTTGCACGCCAATTTTCTTAAAGGAATAATTTAGAAAGCCAAAGTCAGAAACCGG 1448  
1593 TGTCTGCTCTACTTTCCCTGGAGCAAGCGTTGTTTACCAGCTATATCTTCTTCAAGACCCA 1652  
1449 TGTCTGCTCTACTTTCCCTGGAGCAAGCGTTGTTTACCAGCTATATCTTCTTCAAGACCCA 1508

1653 GTGATGCTGCTGATGAAGGTAGAGACACGCTCCATTAAATAGTCCAGAAATGCAACAGATG 1712  
Db GTGATGCTGCTGATGAAGGTAGAGACACGCTCCATTAAATAGTCCAGAAATGCAACAGATG 1568  
1713 AATCCCTACGAAGGAGTTGATTTGCAAAATCTGGCTGAGCATATTTCTATTCACCTGCTAGCA 1772  
Db AATCCCTACGAAGGAGTTGATTTGCAAAATCTGGCTGAGCATATTTCTATTCACCTGCTAGCA 1628  
1773 AGTCTGTGTGCATTTATGTGCACACACATTTGTGGCTTGCCTGCTCTCTACAGACACAGGC 1832  
Db AGTCTGTGTGCATTTATGTGCACACACATTTGTGGCTTGCCTGCTCTCTACAGACACAGGC 1688  
1833 AGGGAATTTGATCTCTCCACATTTGGTTCGAAGACTTTCTTTGTGATGAAGAGAGAACTCTGG 1892  
Db AGGGAATTTGATCTCTCCACATTTGGTTCGAAGACTTTCTTTGTGATGAAGAGAGAACTCTGG 1748  
1893 CTGCTGATTTTGACCTGGGGTTCTCAGGAAATTCAGAAAGATAGTAATGTCATGCCATAC 1952  
Db CTGCTGATTTTGACCTGGGGTTCTCAGGAAATTCAGAAAGATAGTAATGTCATGCCATAC 1808  
1953 AGTCTGTGGGAAATTTGTGCACAAATCACCCACACTAGCAGGAACGATGAGTTTATTCATCA 2012  
Db AGTCTGTGGGAAATTTGTGCACAAATCACCCACACTAGCAGGAACGATGAGTTTATTCATCA 1868  
2013 CCCCCAGCACAACTGTCCCATCAGTCTTTCGAACCTCAACTTCTACAGCAATGGGGTACTTC 2072  
Db CCCCCAGCACAACTGTCCCATCAGTCTTTCGAACCTCAACTTCTACAGCAATGGGGTACTTC 1928  
2073 ATGCTCTTTATCATGAGGCGCATCATAGCTTTGCGACCTTTATGCAAGTTCTGAAACAAGAGG 2132  
Db ATGCTCTTTATCATGAGGCGCATCATAGCTTTGCGACCTTTATGCAAGTTCTGAAACAAGAGG 1988  
2133 GACTGGGGGTCCCACTAGCACCCCACTAACTGTATCATGCCAGGAGCAGCTGTGGGA 2192  
Db GACTGGGGGTCCCACTAGCACCCCACTAACTGTATCATGCCAGGAGCAGCTGTGGGA 2048  
2193 AGCGGGCAGCCTGTGCTACTTCTCTCAATGAAGGCACCATCTCACTGCCCTTGGCAGA 2252  
Db AGCGGGCAGCCTGTGCTACTTCTCTCAATGAAGGCACCATCTCACTGCCCTTGGCAGA 2108  
2253 CATTTTACCAAGTCTGCTATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTCTACAG 2312  
Db CATTTTACCAAGTCTGCTATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTCTACAG 2168  
2313 TGGCAGACACATGACCCAGGAGATATCAGTCTAGTCTTGTGCTGAGCAGCAGTGGACA 2372  
Db TGGCAGACACATGACCCAGGAGATATCAGTCTAGTCTTGTGCTGAGCAGCAGTGGACA 2228  
2373 AGAAGCTTCCAGAACCTTTGCTTTGGAGAAAGTATGATGAAGAAAGATGAAGACAGTACTTG 2432  
Db AGAAGCTTCTGAACTTTGCTTTGGAGAAAGTATGATGAAGAAAGATGAAGACAGTACTTG 2288  
2433 GGGAGGAAAGCAGAGATTTGCTACTGAAGTGAAGCCAAATCCAAAGGAGCACCAGAGTTTA 2492  
Db GGGAGGAAAGCAGAGATTTGCTACTGAAGTGAAGCCAAATCCAAAGGAGCACCAGAGTTTA 2348  
2493 TCACCTTCTTACAGACTCTTGGGCTTTTGTGGGGGCTACAGCTCTGCTGCCATCT 2552  
Db TCACCTTCTTACAGACTCTTGGGCTTTTGTGGGGGCTACAGCTCTGCTGCCATCT 2408  
2553 TTGTTTCAACACTTTCAGTGGTCTCTGTTCCAGAACTGTAGTATCTGCAAAAGTTGCACAAAAT 2612  
Db TTGTTTCAACACTTTCAGTGGTCTCTGTTCCAGAACTGTAGTATCTGCAAAAGTTGCACAAAAT 2468  
2613 ACCTAATAACCAAGAAAGATTTGTCAGTATATGCTGAGAGTGCCACATATTTGTC 2672  
Db ACCTAATAACCAAGAAAGATTTGTCAGTATATGCTGAGAGTGCCACATATTTGTC 2528  
2673 TTGTAAGATGCTGTGAAATCTTTAAGGATATTTGGGTTTTTCAAGGAGACCAACAAA 2732  
Db TTGTAAGATGCTGTGAAATCTTTAAGGATATTTGGGTTTTTCAAGGAGACCAACAAA 2588  
2733 AGAGAGTGTCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAAAAC 2792

2589 AGAGAGTGTCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAAAAC 2648  
QY TTCTCAATATATTTCTGAGTTTTTGTGGTCTGTAGGTAAAGTGTGGCACTGCTGGCAAT 2852  
Db TTCTCAATATATTTCTGAGTTTTTGTGGTCTGTAGGTAAAGTGTGGCACTGCTGGCAAT 2708  
2853 GAAGGTCATCAGATGATGTTCTTGTAGGTACAGACTTCTGGCTCAAGAGTTTGAAGGTGC 2912  
Db GAAGGTCATCAGATGATGTTCTTGTAGGTACAGACTTCTGGCTCAAGAG-TTGAAGGTGC 2767  
2913 CTTTCGAGGGGTGAGGCTTGCCTTGCCTGTCGAGTATCTCTGCGAAGACAGTGCCTTCT 2972  
Db CTTTCGCA-GGGTCAGGCTTGCCTTGCCTGTCGAGTATCTCTGCGAAGACAGTGCCTTCT 2826  
2973 NCCTCCATCGATCTGAGATCTTCCAGCTTT 3003  
Db CCCTCCATCGATCTGATCTTCCAGCTCT 2857

RESULT 6  
AX431391  
LOCUS AX431391 2487 bp DNA linear PAT 28-JUN-2002  
DEFINITION Sequence 24 from Patent WO0240666.  
ACCESSION AX431391  
VERSION AX431391.1 GI:21656238  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Enkins D.K., Winther, M.D., Haardt M., Goldberg, Y.P., Nwaka, S.O.,  
Ponton, A., Allen, S.J., de Antueno, R.J., and Knickle, L.C.  
TITLE Fat regulated genes, uses thereof, and compounds for modulating  
same  
JOURNAL Patent: WO 0240666-A 24 23-MAY-2002;  
XENON GENETICS INC (CA)  
FEATURES  
source  
1. .2487  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 706 a 548 c 570 g 663 t  
ORIGIN  
Query Match 82.8%; Score 2485.4; DB 6; Length 2487;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2486; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 341 ATGGATGAATCTGCACTGACCCCTTGGTACAATAGATGTTTCTTATCTGCCACATTTCATCA 400  
Db 1 ATGGATGAATCTGCACTGACCCCTTGGTACAATAGATGTTTCTTATCTGCCACATTTCATCA 60  
QY 401 GAATACAGTGTGTGTCGATGTAAGCACACAAAGTGAAGAAATGGGGTGAAGTGTGGCTTTTGA 460  
Db 61 GAATACAGTGTGTGTCGATGTAAGCACACAAAGTGAAGAAATGGGGTGAAGTGTGGCTTTTGA 120  
QY 461 CCACCGCTTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAATAGTTCGGAAGAG 520  
Db 121 CCACCATCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAATAGTTCGGAAGAG 180  
QY 521 CCATTTGTTGGAAGATGTTTACTCTGCACCTCCCGAGAGCTGGGACAAAATTTTTCAC 580  
Db 181 CCATTTGTTGGAAGATGTTTACTCTGCACCTCCCGAGAGCTGGGACAAAATTTTTCAC 240  
QY 581 CCAGTATCCCGCTTTTGGGTTTTCGGAATGTTTATTTATATCAATGAAACTTCACAGA 640  
Db 241 CCAGTATCCCGCTTTTGGGTTTTCGGAATGTTTATTTATATCAATGAAACTTCACAGA 300  
QY 641 CACCGGGAATGGCTTGCAGAGCCCTTTCTAGCTTTTATTTATCAAGAGCAGATGTG 700  
Db 301 CACCGGGAATGGCTTGCAGAGCCCTTTCTAGCTTTTATTTATTTCAAGAGCAGATGTG 360



QY 701 CATAGGGGATGTTTGGCCACCAATGTGACTGAAAAATGCTGAAACAGCAGTAGAGTACAA 760  
DB |||||  
361 CATAGGGGATGTTTGGCCACCAATGTGACTGAAAAATGCTGAAACAGCAGTAGAGTACAA 420  
QY 761 GAGGCAATTGGCAGAGTGGCTGCTGAAATTAACCCCTGATGGTTCTGCCCCAGCAGCAATCA 820  
DB |||||  
421 GAGGCAATTGCAGAGTGGCTGCTGAAATTAACCCCTGATGGTTCTGCCCCAGCAGCAATCA 480  
QY 821 AAAGCCGTTAAACAAAGTGAAGAAAGCTAAAGGATTTCTTCAAGAAATGGTTGCCACT 880  
DB |||||  
481 AAAGCCGTTAAACAAAGTGAAGAAAGCTAAAGGATTTCTTCAAGAAATGGTTGCCACT 540  
QY 881 GTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTGCTAAACCTGTTCAACAGCTTC 940  
DB |||||  
541 GTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTGCTAAACCTGTTCAACAGCTTC 600  
QY 941 TTTTGGAAACATTTCAAAATTCACAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACG 1000  
DB |||||  
601 TTTTGGAAACATTTCAAAATTCACAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACG 660  
QY 1001 AATTTCGCGCTTCTGTTCTACCCAGTTCATAGATCCCATATTTGACTATCTGCTGCTCACT 1060  
DB |||||  
661 AATTTCGCGCTTCTGTTCTACCCAGTTCATAGATCCCATATTTGACTATCTGCTGCTCACT 720  
QY 1061 TTCAATCTCTTCTGCCATAAACATCAAGCAACATACATATTTGCTTCAGGCAATTAATCTCAAC 1120  
DB |||||  
721 TTCAATCTCTTCTGCCATAAACATCAAGCAACATACATATTTGCTTCAGGCAATTAATCTCAAC 780  
QY 1121 ATCCCAATCTTCAGTACCTTGATCATTAAGCTTGGGGCTTCTTCATAGCAGGAAGGCTC 1180  
DB |||||  
781 ATCCCAATCTTCAGTACCTTGATCATTAAGCTTGGGGCTTCTTCATAGCAGGAAGGCTC 840  
QY 1181 GATGAAACACAGATGGGCAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1240  
DB |||||  
841 GATGAAACACAGATGGGCAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
QY 1241 GTTGAATTTCTTCGACAGCAATTTCTGGAGATCTTCTGGAAGGACACGTTCTAGG 1300  
DB |||||  
901 GTTGAATTTCTTCGACAGCAATTTCTGGAGATCTTCTGGAAGGACACGTTCTAGG 960  
QY 1301 AGTGGAAAACTCTTGTGCTCGGCAGCACTTTGTGAGTGTGGTAGACTCTGCT 1360  
DB |||||  
961 AGTGGAAAACTCTTGTGCTCGGCAGCACTTTTGTGAGTGTGGTAGACTCTGCT 1020  
QY 1361 ACCAATGTCATCCAGACATCTTGATATACCTGTTGGAATCTCCTATGATCGCATATC 1420  
DB |||||  
1021 ACCAATGTCATCCAGACATCTTGATATACCTGTTGGAATCTCCTATGATCGCATATC 1080  
QY 1421 GAAGTCACTACAATGGTGAACCACTGGGCAAACTTAAGAAATGAGAGCTGTGGAGT 1480  
DB |||||  
1081 GAAGTCACTACAATGGTGAACCACTGGGCAAACTTAAGAAATGAGAGCTGTGGAGT 1140  
QY 1481 GTAGCAAGAGGTGTTATTAGAAATGTTACGAAAACTATGGTTGTTCCGAGTGGATTTT 1540  
DB |||||  
1141 GTAGCAAGAGGTGTTATTAGAAATGTTACGAAAACTATGGTTGTTCCGAGTGGATTTT 1200  
QY 1541 GCACAGCCATTTCTTAAAGGAAATTTAGAAAGCCTAAAGCTCAGAAACCGGTGCTGCT 1600  
DB |||||  
1201 GCACAGCCATTTCTTAAAGGAAATTTAGAAAGCCTAAAGCTCAGAAACCGGTGCTGCT 1260  
QY 1601 CTACTTTCTCCCTGGAGCAAGCGTTGTACCACTACTTCTTCAAGACCCAGTGTGCT 1660  
DB |||||  
1261 CTACTTTCTCCCTGGAGCAAGCGTTGTACCACTACTTCTTCAAGACCCAGTGTGCT 1320  
QY 1661 GCTGATGAAGGTAGAGACAGTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1720  
DB |||||  
1321 GCTGATGAAGGTAGAGACAGTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
QY 1721 CGAAGGAGGTGATTGCAAACTGGCTGAGCATATTTCTTCACTGTCTAGCAAGTCTGCT 1780  
DB |||||  
1381 CGAAGGAGGTGATTGCAAACTGGCTGAGCATATTTCTTCACTGTCTAGCAAGTCTGCT 1440

QY 1781 GCCATTATGTCCACACACATTTGTGGCTTGCTGCTCTCTACAGACACAGCAGGGAAT 1840  
DB |||||  
1441 GCCATTATGTCCACACACATTTGTGGCTTGCTGCTCTCTACAGACACAGCAGGGAAT 1500  
QY 1841 GATCTCTCCATTTGGTCGAGACTTCTTTGTGATGAAGAGGAGTCTCTGGCTGCTGAT 1900  
DB |||||  
1501 GATCTCTCCATTTGGTCGAGACTTCTTTGTGATGAAGAGGAGTCTCTGGCTGCTGAT 1560  
QY 1901 TTTGACCTGGGGTTCTCAGGAAATTCAGAAATGTAGTAATGCATGCCATACAGCTGCTG 1960  
DB |||||  
1561 TTTGACCTGGGGTTCTCAGGAAATTCAGAAATGTAGTAATGCATGCCATACAGCTGCTG 1620  
QY 1961 GGAATTTGTGTCACAATCACCACACTAGCAGGAAACGATGAGTTTTTATCACCCCCAGC 2020  
DB |||||  
1621 GGAATTTGTGTCACAATCACCACACTAGCAGGAAACGATGAGTTTTTATCACCCCCAGC 1680  
QY 2021 ACAAATGTCATCAGTCTTCCGAACTCAACTTCTACAGCAATGGGGTACTTCACTGCTCT 2080  
DB |||||  
1681 ACAAATGTCATCAGTCTTCCGAACTCAACTTCTACAGCAATGGGGTACTTCACTGCTCT 1740  
QY 2081 ATCATGAGGCGCATATAGCTTGCAGCTTTATGCAGTTCTGAAACAAGAGGGACTGGGG 2140  
DB |||||  
1741 ATCATGAGGCGCATATAGCTTGCAGCTTTATGCAGTTCTGAAACAAGAGGGACTGGGG 1800  
QY 2141 GGTCCCACTAGCACCCCACTTAACCTGATCAGCAGGAGCAGCTGGTGCAGGAGCGGCC 2200  
DB |||||  
1801 GGTCCCACTAGCACCCCACTTAACCTGATCAGCAGGAGCAGCTGGTGCAGGAGCGGCC 1860  
QY 2201 AGCTGTGCTACCTTCTCCATGAAGGACCACTCTCAGTGCCTTGGCCAGACATTTTAC 2260  
DB |||||  
1861 AGCTGTGCTACCTTCTCCATGAAGGACCACTCTCAGTGCCTTGGCCAGACATTTTAC 1920  
QY 2261 CAACTGCTGCCATGAAACAGTAGGAAAGTTTATCAAGTATGCAATTTCTTACAGTGGCAG 2320  
DB |||||  
1921 CAACTGCTGCCATGAAACAGTAGGAAAGTTTATCAAGTATGCAATTTCTTACAGTGGCAG 1980  
QY 2321 CACGATCAGCAGGAGATATCAGTCTTAGTCTTCTGAGCAGCAGTGGGCAAGAGCTT 2380  
DB |||||  
1981 CACGATCAGCAGGAGATATCAGTCTTAGTCTTCTGAGCAGCAGTGGGCAAGAGCTT 2040  
QY 2381 CCAGAACCTTTGCTTGGAGAAAGTGAAGAAATGAAGACAGTGAATTTGGGGAGGAA 2440  
DB |||||  
2041 CCAGAACCTTTGCTTGGAGAAAGTGAAGAAATGAAGACAGTGAATTTGGGGAGGAA 2100  
QY 2441 CAGCGAGATTCCTACCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2500  
DB |||||  
2101 CAGCGAGATTCCTACCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2160  
QY 2501 TTACAGAGACTCTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2560  
DB |||||  
2161 TTACAGAGACTCTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2220  
QY 2561 AACTTCAGTGGTCTGTTTCAGAACTGAGTATCTGAAAGTGGCAGAAATACCTAATA 2620  
DB |||||  
2221 AACTTCAGTGGTCTGTTTCAGAACTGAGTATCTGAAAGTGGCAGAAATACCTAATA 2280  
QY 2621 ACCAGAACAGAAAGAAATGTTGAGTATGCTGAGTGGCAGTATGCTTCTGTTGAG 2680  
DB |||||  
2281 ACCAGAACAGAAAGAAATGTTGAGTATGCTGAGTGGCAGTATGCTTCTGTTGAG 2340  
QY 2681 AATGCTGAAATGTTTAAAGGATTTGGGGTTTCAAGGAGACCAACAAAGAGAGTG 2740  
DB |||||  
2341 AATGCTGAAATGTTTAAAGGATTTGGGGTTTCAAGGAGACCAACAAAGAGAGTG 2400  
QY 2741 TCTGTTTAAAGTGAAGCAGTATGCTGAGTGGTCTGAGTGGTCTGAGTGGTCTGAG 2800  
DB |||||  
2401 TCTGTTTAAAGTGAAGCAGTATGCTGAGTGGTCTGAGTGGTCTGAGTGGTCTGAG 2460  
QY 2801 TATATTCTGAGTGGTGGTCTGAG 2827  
DB |||||  
2461 TATATTCTGAGTGGTGGTCTGAG 2487

RESULT 7	AX420455	2484 bp	DNA	linear	PAT 18-JUN-2002
LOCUS	Sequence 3 from Patent WO0216592.				
DEFINITION	AX420455				
ACCESSION	AX420455				
VERSION	AX420455.1	GI:21524604			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				
AUTHORS	Kapeller-Libermann, R. and Logan, T.J.				
TITLE	Atcr-1, a human acyltransferase and uses thereof				
JOURNAL	Patent: WO 0216592-A 3 28-FEB-2002;				
	Millennium Pharmaceuticals, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..2484				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
BASE COUNT	704 a 548 c 570 g 662 t				
ORIGIN					
Query Match	82.7%; Score 2484; DB 6; Length 2484;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 2484;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	341 ATGGATGAATCTGCATGACCTTGGTACAAATAGATGTTTCTATCTGCCACATTCATCA	400			
Db	1 ATGGATGAATCTGCATGACCTTGGTACAAATAGATGTTTCTATCTGCCACATTCATCA	60			
QY	401 GAATACAGTGTGGTGCATGTAAAGCACACAAAGTGGAGTGGGTGAGTGTGGCTTTAGA	460			
Db	61 GAATACAGTGTGGTGCATGTAAAGCACACAAAGTGGAGTGGGTGAGTGTGGCTTTAGA	120			
QY	461 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAAGAAAGCCCTAATAGTTCGGAAGG	520			
Db	121 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAAGAAAGCCCTAATAGTTCGGAAGG	180			
QY	521 CCATTGTTGGAGATGTTGTTACTCTGCACTCCCAGAGCTGGGACAAATTTTCAAC	580			
Db	181 CCATTGTTGGAGATGTTGTTACTCTGCACTCCCAGAGCTGGGACAAATTTTCAAC	240			
QY	581 CCAGTATCCCGTCTTGGGTTGGGAATGTTATATCAATGAACCTCACACAAGA	640			
Db	241 CCAGTATCCCGTCTTGGGTTGGGAATGTTATATCAATGAACCTCACACAAGA	300			
QY	641 CACCGGGATGGCTTGCAGACGCTTTCTTACGTTCTTTTATTCAGAGCGAGATGTG	700			
Db	301 CACCGGGATGGCTTGCAGACGCTTTCTTACGTTCTTTTATTCAGAGCGAGATGTG	360			
QY	701 CATAGGGCATGTTTGGCCAAATGTGACTGAAATGTGCTGAACAGCAGTAGATACAA	760			
Db	361 CATAGGGCATGTTTGGCCAAATGTGACTGAAATGTGCTGAACAGCAGTAGATACAA	420			
QY	761 GAGCAATTCAGAGTGGCTGCTGAATTAACCCCTGATGTTCTGCCAGAGCAATCA	820			
Db	421 GAGCAATTCAGAGTGGCTGCTGAATTAACCCCTGATGTTCTGCCAGAGCAATCA	480			
QY	821 AAAGCCGTTAAACAAAGTGAAGAAAGAGCTTAAAGGATCTTCAAGAAATGGTTGCCACT	880			
Db	481 AAAGCCGTTAAACAAAGTGAAGAAAGAGCTTAAAGGATCTTCAAGAAATGGTTGCCACT	540			
QY	881 GTCTACCGGCATATGATCAGATGACTGGTGGGTGCTGCTAAAACCTGTTCAACAGCTTC	940			
Db	541 GTCTACCGGCATATGATCAGATGACTGGTGGGTGCTGCTAAAACCTGTTCAACAGCTTC	600			
QY	941 TTTTGGAAATTCAAATTCACAAAGTCAACTTTCAGATGTTTAAAGCTGCAACTGAGACG	1000			
Db	601 TTTTGGAAATTCAAATTCACAAAGTCAACTTTCAGATGTTTAAAGCTGCAACTGAGACG	660			
QY	1001 AATTTGCCGCTTCTGTTTCTACAGTTTCATAGATCCCATATTCATCTGCTCACT	1060			

Db	661 AATTTGCCGCTTCTGTTTCTACAGTTTCATAGATCCCATATTCATCTGCTCACT	720			
QY	1061 TTCATTCTCTTCTGCCATAACATCAAGCACCATACATTCGCTTCAGGCAATATCTCAAC	1120			
Db	721 TTCATTCTCTTCTGCCATAACATCAAGCACCATACATTCGCTTCAGGCAATATCTCAAC	780			
QY	1121 ATCCCAATCTTCAGTACCTTTGATCCATAAGCTTTGGGGCTTCTTCATACGACGAAGCTC	1180			
Db	781 ATCCCAATCTTCAGTACCTTTGATCCATAAGCTTTGGGGCTTCTTCATACGACGAAGCTC	840			
QY	1181 GATGAACACCCAGATGAGCGGAAGATGTTCTCTATAGAGCTTTGCTCCTAGGGCATATA	1240			
Db	841 GATGAACACCCAGATGAGCGGAAGATGTTCTCTATAGAGCTTTGCTCCTAGGGCATATA	900			
QY	1241 GTTGAATTAATTCGACAGCAGCAATTCCTTGGAGATCTCCTCGAAGGCAACCTTCTAGG	1300			
Db	901 GTTGAATTAATTCGACAGCAGCAATTCCTTGGAGATCTCCTCGAAGGCAACCTTCTAGG	960			
QY	1301 AGTGAAGAAACCTTCTGCTCGGGCAGGACTTTTGTCTAGTTGTGGTAGATCTCTGTCT	1360			
Db	961 AGTGAAGAAACCTTCTGCTCGGGCAGGACTTTTGTCTAGTTGTGGTAGATCTCTGTCT	1020			
QY	1361 ACCAATGTCTATCCAGACATCTTGATTAATACCTTTGGAAATCTCTATGATCGCATATC	1420			
Db	1021 ACCAATGTCTATCCAGACATCTTGATTAATACCTTTGGAAATCTCTATGATCGCATATC	1080			
QY	1421 GAAGTCACTACAAATGGTGAACAACTGGGCAACCTTAAGAAATGAGAGCTGTGGAGT	1480			
Db	1081 GAAGTCACTACAAATGGTGAACAACTGGGCAACCTTAAGAAATGAGAGCTGTGGAGT	1140			
QY	1481 GTAGCAAGAGTGTATTAGAATGTTTACGAAACAACTATGGTGTCTCGAGTGGATTTT	1540			
Db	1141 GTAGCAAGAGTGTATTAGAATGTTTACGAAACAACTATGGTGTCTCGAGTGGATTTT	1200			
QY	1541 GCAAGCCATTTTCTTAAAGGAATTTTAGAAAGCAAAGTCAGAAAACCGGTGTCTGCT	1600			
Db	1201 GCAAGCCATTTTCTTAAAGGAATTTTAGAAAGCAAAGTCAGAAAACCGGTGTCTGCT	1260			
QY	1601 CTACTTTCCTGGAGCAAGGTTGTACAGCTATATCTTCCCTCAAGACCCAGTGTGCT	1660			
Db	1261 CTACTTTCCTGGAGCAAGGTTGTGTACAGCTATATCTTCCCTCAAGACCCAGTGTGCT	1320			
QY	1661 GCTGATGAAGGTAGACACACGCTCCATTAATAGTCCAGAAATGCAACAGATGAATCCCTA	1720			
Db	1321 GCTGATGAAGGTAGACACACGCTCCATTAATAGTCCAGAAATGCAACAGATGAATCCCTA	1380			
QY	1721 CGAAGAGGTTGATTGCAAAATCTGGCTGAGCATATTTCTATTCCTCTAGCAAGTCTGT	1780			
Db	1381 CGAAGAGGTTGATTGCAAAATCTGGCTGAGCATATTTCTATTCCTCTAGCAAGTCTGT	1440			
QY	1781 GCCATTATGTCACACACATTTGGCTTGCCTGCTCTCTACAGACACAGGCAAGGAAT	1840			
Db	1441 GCCATTATGTCACACACATTTGGCTTGCCTGCTCTCTACAGACACAGGCAAGGAAT	1500			
QY	1841 GATCTCTCACATTTGTCGAAGACTTCTTTGTGATGAAAGAGGAAGTCTCTGGCTCGTAT	1900			
Db	1501 GATCTCTCACATTTGTCGAAGACTTCTTTGTGATGAAAGAGGAAGTCTCTGGCTCGTAT	1560			
QY	1901 TTTGACCTGGGGTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGTCTGT	1960			
Db	1561 TTTGACCTGGGGTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGTCTGT	1620			
QY	1961 GGAAATTTGTTCACATACCCACACTAGCAGGAACGATGATTTTTTATCACCCCAGC	2020			
Db	1621 GGAAATTTGTTCACATACCCACACTAGCAGGAACGATGATTTTTTATCACCCCAGC	1680			
QY	2021 ACAACTGTCCCATCAGTCTTCGAACTCAACTTCTACAGCAATGGGGTACTTCATGTCTTT	2080			
Db	1681 ACAACTGTCCCATCAGTCTTCGAACTCAACTTCTACAGCAATGGGGTACTTCATGTCTTT	1740			
QY	2081 ATCATGGAGGCCATCATAGCTTGCAGCTTTATGCAAGTCTGAACAAAGAGGGAAGCTGGGG	2140			

```
Db 1741 ATCATGGAGGCCATCATAGCTTGGAGCCTTTATGACGTTCTGAACAAGAGGGGACTGGGG 1800
Qy 2141 GGTCCCACTAGCACCCACCTAACTGATACAGCAGGAGCAGCTGGTGGGAAGCGGCC 2200
Db 1801 GGTCCCACTAGCACCCACCTAACTGATACAGCAGGAGCAGCTGGTGGGAAGCGGCC 1860
Qy 2201 AGCCTGTGCTACCTTCTCTCCAAATGAAGGACCATCTACGCTTGGCCCTGGCAGACATTTTAC 2260
Db 1861 AGCCTGTGCTACCTTCTCTCCAAATGAAGGACCATCTACGCTTGGCCCTGGCAGACATTTTAC 1920
Qy 2261 CAAAGTCTGCTCATGAACAAGTAGAAGTTTATCCAGTATGGCATTTTACAGTGGCAGAG 2320
Db 1921 CAAAGTCTGCTCATGAACAAGTAGAAGTTTATCCAGTATGGCATTTTACAGTGGCAGAG 1980
Qy 2321 CACGATGACCAAGGAAGATATACGTCCTAGTCTTGTCTGAGCAGCAGTGGGACAAAGAGCTT 2380
Db 1981 CACGATGACCAAGGAAGATATACGTCCTAGTCTTGTCTGAGCAGCAGTGGGACAAAGAGCTT 2040
Qy 2381 CCAGAACCTTTGCTCTTGGAAGATGATGAAGAAGATGAAGACAGTGAATTTGGGGAGGAA 2440
Db 2041 CCAGAACCTTTGCTCTTGGAAGATGATGAAGAAGATGAAGACAGTGAATTTGGGGAGGAA 2100
Qy 2441 CAGCGAGATTTGCTACCTGAAGTGAAGCAATCCAGGAGCAGCAGCAGTTCACCTTC 2500
Db 2101 CAGCGAGATTTGCTACCTGAAGTGAAGCAATCCAGGAGCAGCAGCAGTTCACCTTC 2160
Qy 2501 TTACAGAGACTCTTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2560
Db 2161 TTACAGAGACTCTTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2220
Qy 2561 AACTTCAGTGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAAATACCTAATA 2620
Db 2221 AACTTCAGTGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAAATACCTAATA 2280
Qy 2621 ACCGAAACAGAAAGAAATGTTGCAATATCTGAGAGTGCACATATGTTCTGTGAAG 2680
Db 2281 ACCGAAACAGAAAGAAATGTTGCAATATCTGAGAGTGCACATATGTTCTGTGAAG 2340
Qy 2681 AATGCTGTGAAAATCTTTAAGGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTG 2740
Db 2341 AATGCTGTGAAAATCTTTAAGGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTG 2400
Qy 2741 TCTGTTTATAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAAACTTCTAGAA 2800
Db 2401 TCTGTTTATAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAAACTTCTAGAA 2460
Qy 2801 TATATCTGAGTTTGTGGTGTGCTG 2824
Db 2461 TATATCTGAGTTTGTGGTGTGCTG 2484
```

```
RESULT 8
HSM803772 HSM803772 5689 bp mRNA linear PRI 13-MAY-2003
LOCUS Homo sapiens mRNA; cDNA DKFp313H071 (from clone DKFp313H071).
DEFINITION AL832464
ACCESSION AL832464.1 GI:21733030
VERSION 1 (bases 1 to 5689)
KEYWORDS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5689)
AUTHORS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
```

German Genome Project.  
This clone (DKFp313H071) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
information about the clone and the sequencing project is available  
at <http://mips.gsf.de/proj/cDNA/>.

## FEATURES

Location/Qualifiers  
1..5689  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFp313H071"  
/tissue\_type="human fetal skin"  
/clone\_lib="313 (synonym: hlcc2). Vector pTriplEx2; host  
DH10B; sites SfiIA + SfiIB"  
/dev\_stage="fetal"  
polyA\_site 5654  
BASE COUNT 1521 a 1149 c 1256 g 1763 t  
ORIGIN

Query Match 74.7%; Score 2241.8; DB 9; Length 5689;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2246; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Qy 211 CAGCAGCAGCGGCTCCCTGTTGTATGGACATTTCTGCACCCGAACTGATAGCTGAGTC 270  
Db 37 CAGCAGCAGCGGCTCCCTGTTGTATGGACATTTCTGCACCCGAACTGATAGCTGAGTC 96  
Qy 271 CTGAAGTTTATGTTATGAAACAGAAAGACTTTTCATCCAGCACATGATTTGGGAATTAC 330  
Db 97 CTGAAGTTTATGTTATGAAACAGAAAGACTTTTCATCCAGCACATGATTTGGGAATTAC 156  
Qy 331 ACTTTGTGACATGATGAATCTGCACCTGACCTTGGTACATAGATGTTTCTTATCTGCC 390  
Db 157 ACTTTGTGACATGATGAATCTGCACCTGACCTTGGTACATAGATGTTTCTTATCTGCC 216  
Qy 391 ACATTTCATCAATACAGTGTGTTGTCGATGTAAGCACACACAGTGAAGGAAATGGGTGAGTG 450  
Db 217 ACATTTCATCAATACAGTGTGTTGTCGATGTAAGCACACACAGTGAAGGAAATGGGTGAGTG 276  
Qy 451 TGGCTTTAGACCCACCGCTTTCAGATCTGAACTTTTAAATGGAAGAACCCCTAAATGAG 510  
Db 277 TGGCTTTAGACCCACCGCTTTCAGATCTGAACTTTTAAATGGAAGAACCCCTAAATGAG 336  
Qy 511 TCGAAAAAGGCCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGACAGCTCGGACAA 570  
Db 337 TCGAAAAAGGCCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGACAGCTCGGACAA 396  
Qy 571 ATTTTTCACACCCCAAGTATCCCGTCTTTGGGTTTCGGGAATGTTATTTATATCAATGAAC 630  
Db 397 ATTTTTCACACCCCAAGTATCCCGTCTTTGGGTTTCGGGAATGTTATTTATATCAATGAAC 456  
Qy 631 TCACACAGACACCGCGGATGGCTTTCAGAACGCTTTCCTTACGTTCTTTTATTTCAAGA 690  
Db 457 TCACACAGACACCGCGGATGGCTTTCAGAACGCTTTCCTTACGTTCTTTTATTTCAAGA 516  
Qy 691 GCGAGATGTGCATAAAGGCGCATTTTGGCCACCAATGTGACTGAAAATGTGCTGGAACAGAG 750  
Db 517 GCGAGATGTGCATAAAGGCGCATTTTGGCCACCAATGTGACTGAAAATGTGCTGGAACAGAG 576  
Qy 751 TAGAGTCAAGAGGCAATTTGAGAAGTGGCTGCTGGAATTTAAACCTGATGTTCTTGCCCA 810  
Db 577 TAGAGTCAAGAGGCAATTTGAGAAGTGGCTGCTGGAATTTAAACCTGATGTTCTTGCCCA 636  
Qy 811 GCAGCAATCAAAAGCGCTTAAACAAAGTGAAGAAAGAGCTTAAAGAGATTTCTCAAGAAAT 870  
Db 637 GCAGCAATCAAAAGCGCTTAAACAAAGTGAAGAAAGAGCTTAAAGAGATTTCTCAAGAAAT 696  
Qy 871 GGTTCGCACTGTCTCAGCGGCAATGATCAGACTGCTGGGTGGGTGCTGCTTAAACCTGTT 930  
Db 697 GGTTCGCACTGTCTCAGCGGCAATGATCAGACTGCTGGGTGGGTGCTGCTTAAACCTGTT 756  
Qy 931 CAACAGCTTCTTTTGGAAACATTCAAAATTCACAAAGGTCACACTTGAGATGGTTAAAGCTGC 990





Db	1861	AAGGATATTGGGGTTTTCAGGAGACCAACAAAGAGAGTGCTGTGTTTAGAACTGAGC	1920
Qy	2759	AGCATTCTTCTACTCAATCGACCGACCAAACTCTAGATATATCTGAGTTTGTG	2818
Db	1921	AGCATTCTTCTACTCAATCGACCGACCAAACTCTAGATATATCTGAGTTTGTG	1980
Qy	2819	GTGCTGTAGGTAAAGTGTGCACTGCTGCGCAATGAAGGTCAATGAGTTCCTTGTA	2878
Db	1981	GTGCTGTAGGTAAAGTGTGCACTGCTGCGCAATGAAGGTCAATGAGTTCCTTGTA	2040
Qy	2879	GGTACCAAGTCTTGGCTCAAGAGTTGAAGGTGCTTCGAGGGTCAGCCCTGCTGT	2938
Db	2041	GGTACCAAGTCTTGGCTCAAGAGTTGAAGGTGCTTCGAGGGTCAGCCCTGCTGT	2098
Qy	2939	NCCGAGTGTATCTCCCTGGAGAGCAAGTCTTCCTTCATGATCTGATCTCCCA	2998
Db	2099	CCGAGTGTATCTCCCTGGAGAGCAAGTCTTCCTTCATGATCTGATCTCCCA	2158
Qy	2999	GCTTT 3003	
Db	2159	GCTCT 2163	
RESULT 10			
LOCUS	BC019201	3857 bp mRNA linear	ROD 20-SEP-2002
DEFINITION	Mus musculus, glycerol-3-phosphate acyltransferase, mitochondrial, clone MGC:29133 IMAGE:5041120, mRNA, complete cds.		
ACCESSION	BC019201		
VERSION	1		
KEYWORDS	MGC.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 3857)		
TITLE	Strausberg, R.		
JOURNAL	Direct Submission		
REMARK	Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
COMMENT	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a> Contact: MGC help desk Email: <a href="mailto:cgaps@mail.nih.gov">cgaps@mail.nih.gov</a> Tissue Procurement: Jeffrey E. Green, M.D. cdNA Library Preparation: Life Technologies, Inc. cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a> Contact: <a href="mailto:nisc.mc@nhgri.nih.gov">nisc.mc@nhgri.nih.gov</a> Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maekeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAC Plate: 39 Row: k Column: 10 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6680056. Location/Qualifiers 1..3857 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090"		

CDS	/clone="MGC:29133 IMAGE:5041120" /tissue type="Kidney, normal. 5 month old male mouse." /clone_lib="NCI CGAP_Kid14" /lab_host="DH105" /note="Vector: pCMV-SPORT6" 160..2643 /codon_start=1 /product="glycerol-3-phosphate acyltransferase, mitochondrial" /protein_id="AAH19201.1" /db_xref="GI:17512495" /db_xref="LocustID:14732" /translation="MEESVTVGTITDVSLPSSBSYSLGRKHTSDWDCGPKFTFF GWLARLSVLEQVRDVKHGMFATVTENVLSSRSRQVQAI AEVAELNPDGSAQOQS KAIQVKRKARKILQEMVATVSPGMIRLTGWLKLFNSFFWNIQIHKGQLEVMKAAT ETNLPULFELVHRSHIDYLLTFLFCHNIKAPYIAGNNLNI PVSTLIHLKLGFFI RRDETDPDRKDIYLRALLHGVHVELLQQOFLFLEFLEGRSGKVSARGLLSV VVDTLSSNTIPDILVIVPGISYDRIIEGHNGEQLKPKNESLMSVARGVIRMLRKN YGVVRVDFAPQPSLKEYLGQSKPVSAPLSLEQALLPAILPSRPNDADEHLSLVN ESRNPADEAFRELRILANAEHILFTASKSCAIMSTHIVACLLLYRHRQGHLSLVED FVMKEVLARDFLPGNSNEDVMHAIQLLGNCTVITHTSRKDEFFTPSTVPSV FELNYSNGVLHVIMEALIASIYAVLNKRCGSGSAGLNLISOELVRKAASLCY LLSNEGTTSLPQCTFYQCHETGKFIQYIGILTVAEQDDQSDVSPGLAQWKKLPE LNRWDEDEDDFGEQKDYKVSQSKHQOQFIFLQRLGLPLLEAYSSAAIFVHN FSGPPESEYLOKLHRYLITRTRNVAVVAESATYCLVNKNAVMFKDIDGVFKETQKR VSVLSLSTFLLPQCRQKLLVLSFVVL" BASE COUNT 941 a 952 c 1022 g 942 t ORIGIN		
	Query Match	63.9%; Score 1917.6; DB 10; Length 3857;	
	Best Local Similarity	83.0%; Pred. No. 0;	
	Matches 2235; Conservative	0; Mismatches 446; Indels 11; Gaps 4;	
Qy	309	CAGCACATGTTGGGAATTACACTTTGTGACATGATGATCAATCTGCACTGACCTTGGTA	368
Db	128	CAGCACATGTTGGGACTTGCACGTTCTGCCATGAGGAGTCTTCAGTCACACTGGCA	187
Qy	369	CAATAGATCTTCTTATCTGCCACATTCATCAGATACAGATGTTGGTCAGTGAACACA	428
Db	188	CAATAGACGTTTCTTATCTGCCACGTTTCATCGAATACAGCCTTGGCCCATGTAACACA	247
Qy	429	CAAGTGAGGAATGGGGTGAAGTGTGGCTTTAGACCCACCGCTCTTCAGATCTGCAACTTAA	488
Db	248	CCAGTGAGGACTGGGTGTGACTGTGGGTTCAACACTTACCTTCTTCAGATCTGCAACTGA	307
Qy	489	AATGGAAGAAAGCCCTTAATGAGTCGGAAGAGCCATTTGTTGGAAGATGTTGTATCTCT	548
Db	308	AATGGAAGGAGAGCCTTATGAGCCGAGAGAGGCCCTTCGTGGGAGAGTGTCTATCTCT	367
Qy	549	GCACCTCCCGAGAGCTGGGACAAATTTTCAACCCAGATATCCCGTCTTTGGGTTTCGGA	608
Db	368	GCACGCCACAGAGCTGGGAAAGGTTTTTCAACCCAGATATCCCATCTCTGGGGTTTCGGA	427
Qy	609	ATGTTATTATATCAATGAAATTCACACAAAGACACCGCGGATGGCTTGCAGAACGCTTT	668
Db	428	ATGTTATTATATCAATGAAAGCCACACAAAGGACAGAGATGGCTGGCGAGAGCGCTGT	487
Qy	669	CTTACGCTCTTTTATTCAGAGCGGAGATGTGCAHAAGGGCATGTTTTCACCAATGTGA	728
Db	488	CTTACATCTCTTTTGTTCAGAGCGAGAGCTCCATAAGGGCATGTTTTCACCACTGTTA	547
Qy	729	CTGAAATCTGTGACAGCATGTAGATACAGAGGCAATTCAGAAAGTGGCTGCTGAAT	788
Db	548	CTGAGAATGTACTGACAGCAGCAGAGTCCAGAGAGCAATTCGTGAAGTGGCTGCGGAGT	607
Qy	789	TAAACCTCTGATGGTCTGCGCCAGCAGCAATCAAAAGCCGTAAACAAAGTCAAAAAGAAAG	848
Db	608	TGACCCAGATGATCTGCCAGCAGCAGCTCCAAAGCCATCCAGAGGTGAAAGGAAG	667
Qy	849	CTAAAGAGATCTTCAAGAAATGGTTGGCAGTGTCTTCAACCGGCAATGATCAGACTGACTG	908
Db	668	CCAGGAAGATCTCTCCAGGAGATGGTTCGCCACCGTCTCTCCCGAGGATGATCAGGCTGACTG	727



Qy	909	GGTGGGTGCTGCTAAAAC	TGTTCAACAGCTTC	TTTCTTTTGGAA	CATTTCAAATTC	CAAAAGGTC	968	
Db	728	GCTGGGTGTTACTAAAG	CTCTTCAACAGCTTC	TTCTCTGGAA	CATTCAGATTC	CACAAGGCT	787	
Qy	969	AACTTGAGATGGTTTAAAG	CTGCAACTGAGACGAAT	TTTGC	GGCTTCTGTTTCTAC	CAAGTTC	1028	
Db	788	AACTCGAGATGGTCAAG	GGCTGCACTGAGACGA	AACTTGC	CGCTTGTGTTCTG	CGGTGC	847	
Qy	1029	ATAGATCCCATATTGAC	TATCTGCTGCTCACT	TTTTCATTTCTTCTG	CCATAACATCA	AAAG	1088	
Db	848	ACAGATCCCATTTGAC	TACTACGTGTGCTAC	CTTTGCTACCTT	CATCTTTTGGCA	CAACATCAAG	907	
Qy	1089	CACCATACATTTGCTT	CAGGCAATATCTCA	CACTCCCAATTTCT	CAGTACCTTGN	TCATTA	1148	
Db	908	CGCCGTACATCGCCT	TCGGCAATTAATCT	CAACATCCCCG	CTTCTCAGTAC	CTTGNATTCACA	967	
Qy	1149	AGCTTGGGGGCTTCTT	CATACGACGAAGG	CTCGATGAACACCA	CAGATGACGGA	AAAGATG	1208	
Db	968	AGCTTGGGGGCTTTT	CATAAGACGAGG	CTCGATGAACACCA	CAGATGACGGA	AAAGATG	1027	
Qy	1209	TTCTCTATAGAGCT	TTGCTCATCGGG	CATATAGTTGA	ATTACTTCGAC	CAGCAATTTCT	1268	
Db	1028	TTCTGTACAGAGC	GTGCTCCATGGG	CATGTAGTTGA	CTCTCCGAC	CAGCAGTTCC	1087	
Qy	1269	TGGAGATCTTCTCG	GAAGGACACGTTCT	AGGAGTGGAAAA	CACTCTTGTG	CTCGGCAG	1328	
Db	1088	TGGAGATCTTCTCG	GAAGGACACCGCT	CTCCGACGTGG	CAAGACCTCTG	CTGCGCGGCAG	1147	
Qy	1329	GACTTTTGTGAGTT	GTGTGATGACTCTG	CTACCAATGT	CATCCACAG	ACATCTTGATAA	1388	
Db	1148	GGCTCTCTCAGT	GATGTGATACTCTG	CTGCTCAACA	CCATCCC	CGACATCTCTGTCA	1207	
Qy	1389	TACTGTTTGGAA	TCTCCTATGATCG	CAATATCGA	AGTCTACTCAAT	TGTTGAACA	CTGG	1448
Db	1208	TACCGTGGGCAT	CTCTATGATCG	CATAATCGA	AGGTCTACTCAAT	TGTTGAACA	CTGG	1267
Qy	1449	GCAAACTTAAAGAA	ATGAGAGCTGTG	GAGTGTAGCA	AGGTGTTAT	TAGATGTTAC	1508	
Db	1268	GAAAGCCCAAGAA	GCAGAGCCTTGG	AGTGTGGG	CAGAGGCGTTAT	CAGAAATGCTG	1327	
Qy	1509	GAAAAA	ACTATGTTGTCCG	AGTGGATTTTG	CACAGCCATTTT	CTTAAAGGAATAT	1568	
Db	1328	GAAAAA	ACTACGGCTAC	GTCCGAGTGGAT	TTTGACAGCCAT	TTTCTTGAGGAATAT	1387	
Qy	1569	TAGAAAGCCAAAG	TCAGAAACCGGTG	CTGCTCTATTTCC	TGAGCAAGCGGT	TGTTAC	1628	
Db	1388	TAGAAGCCCAAGT	TCAGAAACCTGTAT	CTGCCCCCTTTCT	CTGGACCAAG	CACTGTTAC	1447	
Qy	1629	CAGCTACTTTCTT	CAAGACCCAGTC	ATGCTGCTGAT	GAGGTAGAGAC	CACCTCCATTA	1688	
Db	1448	CAGCGATCCTTCTT	CAAGACCGAATGAT	GTGCTGATGA	AACTAAGA	CTTATCCAGTA	1507	
Qy	1689	ATGAGTCCGAAAT	TCAACAGATGAA	TCCCTTACGA	AGGAGTTGAT	TGCAAAATCTGGCTG	1748	
Db	1508	ACGAGTCCGAA	AAACCCAGCAGACGA	AGCCTTCCGACGA	AGGCTGAT	TGCAAACTTGGCTG	1567	
Qy	1749	AGCATATTTCTAT	CTGCTAGCAAG	TCGTGTGCCAT	TATGTCCACA	CAATTTGGCTT	1808	
Db	1568	AGCATATTTCTCTT	CACGCCAGCAAG	TCCTGCGCTAT	CATGTCTCAC	CCACATTTGGCTT	1627	
Qy	1809	GCCTGCTCTCT	CTACAGACAGG	CAGGGAATGAT	CTCTCCACAT	TGTTGTCGAGACTTCT	1868	
Db	1628	GTCTGCTCTCTCT	CTACAGACAGG	CAGGGAATCCAT	CTCTCACGCT	TGTTGGAAGACTTCT	1687	
Qy	1869	TTGTGATGAAG	AGGAAAGTCTTG	CTCGTGTGATTTT	TGACCTGGG	GTCTCTCAGGAAATTCAG	1928	
Db	1688	TTTGTATGAAG	GAGGAAGTCTT	AGCTCGCATTT	TCGACCTAGG	CTTCTCCGSGAATTCAG	1747	
Qy	1929	AAGATGTAGT	ATATGCAATGCC	ATCAGCTGCTGG	GAATTTGTGT	CAATCAACCCACACTA	1988	
Db	1748	AAGATGTGT	CTCATGCTAT	TTCAGCTTCTGG	GGAATCTGTGT	CAATCAACCCACACGA	1807	

RESULT 11	
MUSG3PAT	
LOCUS	
DEFINITION	

MUSG3PAT 6634 bp mRNA linear ROD 27-APR-1993  
Mouse glycerol-3-phosphate acyltransferase mRNA, complete cds.



Db 2098 AAACTACGGCTACGTCAGTGGATTTTGCACAGCCATTTTCTTGAAGGAATTTAG 2157  
Qy 1572 AAAGCAAGTCAAGAACCGGTGCTGCTCTACTTTCCCTGGAGCAACGCTTTTACCAG 1631  
Db 2158 AAGGCCAGTCAAGAACCTGTATCTGCCCTTTCTCTGGAGCAACGCTTTTACCAG 2217  
Qy 1632 CTATACCTTCCCTCAAGACCCAGTGTGCTGATGAAGGTAGAGACACGCTCCATTAATG 1691  
Db 2218 CGATCTTCTTCAAGACCCGAATGATGTGCTGATGAACATCAAGACCTATCCAGTAAG 2277  
Qy 1692 AGTCAGAAATGCAACAGATGAATCCCTACGAAGGAGGTGATTCGAATCTGGCTGAGC 1751  
Db 2278 AGTCAGAAATCCAGCAGACGAGCTTCCGACGAGGCTGATTCGAACCTGGCTGAGC 2337  
Qy 1752 ATATTCTATTCACTGCTAGCAAGTCTCTGTCATTTATGCCATTATGCCACACACATTTGGCTTCCC 1811  
Db 2338 ACATTCTCTTACGCCAGCAAGTCTCTGCTATCATGTCCACCCACACATTTGGCTGTC 2397  
Qy 1812 TGCTCTCTACAGACACAGGAGGAATGATCTCTCAATGCTGGAAGTTCGAACTCTTTG 1871  
Db 2398 TGCTCTCTACAGACACAGGAGGAATCCATCTCTCCACGCTTGTGGAAGACTTCTTTG 2457  
Qy 1872 TGATGAAGAGAGTCTCTGCTGCTGATTTTGAACCTGGGTCTTCAGGAATTCAGAG 1931  
Db 2458 TGATGAAGAGAGTCTCTGCTGCTGATTTTGAACCTGGGTCTTCAGGAATTCAGAG 2517  
Qy 1932 ATGTAGTAATGCATGCCATACAGTCTGCTGGAAATTTGTGCACATCACCCACACTAGCA 1991  
Db 2518 ATGTCTCATGATGCTATTACGCTTCTGGGAACTGTGTCAATCACCCACAGGACA 2577  
Qy 1992 GGAACGATGATTTTATACCCGCCAGCAAACTGTCCCATGCTTTCGAACTCAACT 2051  
Db 2578 GGAAGATGATTTTATTAATCTCCAGCAAACTGTCCCGTCACTTTTGAACCTCAACT 2637  
Qy 2052 TCTACAGCAATGGGTACTTCACTGCTTATCATGGAGGCCATCATAGTTCGACCTTT 2111  
Db 2638 TCTACAGCAATGGGTACTTCACTGCTTATCATGGAGGCCATCATAGTTCGACCTTT 2697  
Qy 2112 ATGCAGTCTTGAACAAAGAGGGGACTTGGGGGTCCACCTAGCACCCACCTAACCTGATCA 2171  
Db 2698 ATGCAGTCTTGAACAAAGAGGTGCTCTGGAGGTCCGCTGGAGGCTCGGCAACCTGATCA 2757  
Qy 2172 GCCAGGACGCTGTGTCGGAAGGCGGCGGCTGTGCTACTTCTCTCCAAATGAAGCA 2231  
Db 2758 GCCAGGACGCTGTGTCGGAAGGCGGCGGCTGTGCTACTTCTCTAACGAAGTA 2817  
Qy 2232 CCATCTCACTGCTTGCAGACATTTTACCAAGTCTGACATGAACAGACAGTGAAGTTA 2291  
Db 2818 CCATCTCTGCTTGCAGACATTTTACCAAGTCTGACATGAACAGTTCGACAGTTC 2877  
Qy 2292 TCCAGTATGGCATTTTACAGTGGCAGACGATGACAGGAAGATATCAGTCTAGTC 2351  
Db 2878 TCCAGTATGGCATTTTACAGTGGCAGACGATGACAGGAAGATATCAGTCTAGTC 2937  
Qy 2352 TTGCTGACGACGCTGGGACAAAGCTTCCAGAACCTTTGCTTGGAGAAAGTATGAAG 2411  
Db 2938 TTGCAGAACACGCTGGGACAAAGCTTCCGGAAC---TGAACCTGGAGAACTGACGAG 2994  
Qy 2412 AAGATGAAGACGATGACTTTGGGAGGAAACAGCGAGATGTCTACTGAGGTGAGCCAT 2471  
Db 2995 AAGATGAAGACGATGACTTTGGGAGGAGGAGCGAGATGTCTACTCAAGGTGAGCCAGT 3054  
Qy 2472 CCAAGGAGCACGACGATTTATCACTTCTTACAGAGACTCTTGGGCTTTGCTGGAG 2531  
Db 3055 CCAAGGAGCACGAGAAATTCATCACTTCTTACAGAGCTTCTAGGTCCCTGCTAGAG 3114  
Qy 2532 CCTACGCTCTGCTGCCATCTTTGTTCAAACTTCAAGTGTCTGTTCCAGAACCTGAGT 2591  
Db 3115 CCTACGCTCTGCTGCCATCTTTGTTCAAACTTCAAGTGTCTGTTCCAGAACCTGAGT 3174  
Qy 2592 ATCTGCAAAAGTTGCAAAATACCTAATAACAGAAAGAAATGTTGCACTATG 2651  
Db 3175 ACCTGCAAAAGTTGCAAAATACCTAATAACAGAAAGAAATGTTGCACTATG 3234

Qy 2652 CTGAGAGTGCACATATTGCTTGTGAAGATGCTGTGAATGTTTAAAGGATATTGGG 2711  
Db 3235 CTGAGAGTGCACATATTGCTTGTGAAGATGCTGTGAATGTTTAAAGGATATTGGG 3294  
Qy 2712 TTTTCAAGGAGACCAAAACAAAGAGAGTGTCTGTTTAAAGCTGAGCAGCACTTTTCTAC 2771  
Db 3295 TTTTCAAGGAGACCAAAACAAAGAGAGTGTCTGTTTAAAGCTGAGCAGCACTTTTCTAC 3354  
Qy 2772 CTCAATGCAACCGACAAAACCTTCTAGAAATATATTCTGAGTTTGTGGTGTAGGTAA 2831  
Db 3355 CTCAATGCAACCGACAAAACCTTCTAGAAATATATTCTGAGTTTGTGGTGTAGGTAA 3414  
Qy 2832 CGTGTGACCTGCTGGCAAAATGAAGGTATGAGATGATTCCTTGTAGTACCAGCTTCT 2891  
Db 3415 CTCCT--CAGCACCTATCTCGAGCGAGAGCTGATGATGCTTTCGACCTCCAGTCTGA 3472  
Qy 2892 GGCTCAAGAGTTTGAAGGTGCTTTCGACAGGAGTTCAGGCTGCTGCTGCTGCTGCTGCT 2951  
Db 3473 CCCAGA---GTGCGAGGTATCTTGGCATGCCCA-GCCTGCCCTACCCCGAGGAGACTT 3527  
Qy 2952 CCTGGAAGACAGTGCCTTCTNCCCTCATGATCTGAGATCTTCCACG 3000  
Db 3528 CCAGAAAGACAGTGCCTTCTGCCCCGTGACCTGCGCTCTCCCAAC 3576

## RESULT 12

AX401972 LOCUS AX401972 2646 bp DNA linear PAT 06-JUN-2002  
DEFINITION Sequence 1648 from Patent WO0210453.  
ACCESSION AX401972  
VERSION AX401972.1 GI:21338152  
KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and Elashoff, M.R.  
AUTHORS Molecular toxicology modeling  
TITLE Patent: WO 0210453-A 1648 07-FEB-2002;  
JOURNAL Gene Logic, Inc. (US)  
FEATURES Location/Qualifiers  
source  
1..2646  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/note="EMBL/GenBank Accession No. NM\_017274"

BASE COUNT 670 a 689 c 670 g 617 t  
ORIGIN

Query Match 63.1%; Score 1896.2; DB 6; Length 2646;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2126; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 319 TTTGGGAATTACATTTTGTGACATGATGATCTGCACTGACCCCTTGGTACAATAGATGT 378  
Db 138 TTTGGGACTTCCAGTTGAGTTCATGGAGGAGTCTTCACTGACAAATGGCAATAGACGT 197  
Qy 379 TTCTTATCTGCCAATTCATCAGAAATACAGTTTGGTTCGATGTAAGCACACAACTGAGGA 438  
Db 198 TTCTTATCTGCCAATTCATCAGAAATACAGTTTGGCCGATGTAACACAGAAATAGGA 257  
Qy 439 ATGGGTGAGTGTGGCTTTTAGACCCACCGCTTTCAGATCTGCAACTTTAAATGGAAGA 498  
Db 258 CTGGGTTGACTGTGGCTTCAAACTTCTTTCAGATCCGCAACGCTGAAATGGAAGA 317  
Qy 499 AAGCCTAATGAGTCGGAAGAGGCCATTTGTTGGAAGATGTTTACTCTCTGCACTCCCA 558  
Db 318 GAGCCTCATAGCCGGAAGAGGCCCTTCTGTTGGGAAGGTTGCTTATTCTATCATGACGCTCA 377

Qy	559	GAGCTGGGACAAATTTTTTCAACCCAGTATCCCGCTCTTTGGGTTTGGGAATGTTATTTA	618
Db	378	GAGCTGGGAAGGTTTTTTCAACCCAGTATCCCATCTCTGGGTTTGGGAATGTTATTTA	437
Qy	619	TATCAATGAACACTCACACAGACACCGGATGGCTTGCAGAGCCCTTCTTTACGTTCT	678
Db	438	TATCAATGAAACTCACAAAGGCACCGAGATGGCTGGCAAGCGGCTTTCTTACATCTCT	497
Qy	679	TTTTTATTTCAAGAGCGAGATGTGCATAAAGGGCATGTTTGGCCACCAATGTGACTGAAAAATG	738
Db	498	TTTTTGTTCAGAGGCGGATGTCTCACAGGGCATGTTTGGCCACCAATGACTGACATGT	557
Qy	739	GCTGAAACAGCAGTAGATACAAAGAGGCAATTGCAAGAGTGGCTGCTGAAATTAACACCTGA	798
Db	558	ACTGAATAGCAGCAGATCCAAAGAGGCAATTGCTGAAAGTGGCTGCAGAAATTGAACCCGGA	617
Qy	799	TGGTTCTGCCAGCAGCAATCAAAAGCGGTTAACAAAGTGAAGAAAGAAAGCTAAAGGAT	858
Db	618	TGGATCTGCCAGCAGCAGTCCAAAGGCATCCAGAAAGTGAAGAAAGAAAGCCAGGAAGAT	677
Qy	859	TCTTCAAGAAATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGACTCGGTGGGTGCT	918
Db	678	CCTCCAGGAATGGTTGTCACAGTCTCCCGGGATGATCAGGCTGACTGGCTGGGTGTT	737
Qy	919	GCTAAAACTGTTCACAGCTCTTTTGGAACTTCATAAATTCAAAAGGTCAACTTGAGAT	978
Db	738	ACTAAAGCTCTTCAACAGCTTCTTCTGGAACATTTAGATTCACAAAGGTCAACTTGAGAT	797
Qy	979	GGTTAAAGCTGCAACTCAGAGCAATTTGCGGCTTCTGTTCTACCAAGTTCATAGATCCCA	1038
Db	798	GGTGAAGCTGCAACTGAGAGCAATCTGCGGCTCTTGTCTGCGGTGCACAGATCCCA	857
Qy	1039	TATTGACTATCTGCTGCTCACTTTCATTCTCTCTGCCATAAATCAAAAGCACCATACAT	1098
Db	858	CATCGACTACTGCTGCTCACTTCTCTCTCTGCGCAAAATCAAGCTCCATACAT	917
Qy	1099	TGCTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTGATCATTAAGCTTGGGG	1158
Db	918	CGCTCTGGGCAACAACTCAACATCCCATCTTTCAGTACCTTGATTCACAAGCTTGGGG	977
Qy	1159	CTTCTTCATAGCAGAAAGCTCGATGAAACACAGATGGACGGAAGATGTTCTCTATAG	1218
Db	978	CTTTTTCATAGACGGAGGCTTGAGAACTCTCCAGATGGACGCAAGACATTTCTGTACAG	1037
Qy	1219	AGCTTTGCTCCATGGGATATAGTTGAATTAATCTCGACAGCAGCAATCTTGAGATCTT	1278
Db	1038	AGCTTGCTCCATGGGATATAGTTGAATCTCTCCAGACAGCAGATTTCTTGAGATCTT	1097
Qy	1279	CCTGGAAGGCACAGTTCTAGGAGTGAAAAACCTCTTGTGCTCGGCGAGACTTTTGTCT	1338
Db	1098	CCTGGAAGGCACCGCTCCGCGAGTGGCAGACCTCTGTGTGCGCGGCGGGCTCCTGTC	1157
Qy	1339	AGTTGTTGATAGATACTCTGTCTACAAATGTCATCCAGACATCTTGATATAATCTGTTGG	1398
Db	1158	AGTGGTAGTGATACTCTGTTCATCCAAACCATCCCTGACATCTGCTGTCATCCCTGTGGG	1217
Qy	1399	AATCTCTATGATCGCATATTCGAAGGTCACTACAATGGTGAACAACTGGGCACACCTAA	1458
Db	1218	CATCTGATGATCGGATAATTCGAAGGTCACTACAATGGTGAACAGTTGGGCAAGCCCA	1277
Qy	1459	GAAGAAATGAGACCTGTGGAGTGTAGCAAGAGGTGTTATTAGAAATTTACGAAAAAACTA	1518
Db	1278	GAAGAAATGAAGTCTCTGGAGTGTGGCAGAGGGGTTATCNGAATGCTGCGGAANAATA	1337
Qy	1519	TGTTGTGTCGAGTGGATTTTGCACAGCCATTTTCTTAAAGGAATATTTAGAAAGCCA	1578
Db	1338	CGGCTATGTCCGAGTGGACTTGCACAGCCATTTTCTTTGAAGGAATATTTAGAAAGCCA	1397
Qy	1579	AAGTCAGAAACCGGTGCTGCTACTTCTTCCCTGGACAGCGTGTGATCCAGCTATACT	1638
Db	1398	AAGTCAGAAACCTGTATCTGCTCCCTCTCTTTTGGAGCAAGCACTGTTTACAGCAATCTCT	1457
Qy	1639	TCCTTCAAGACCCAGTGATGCTGTGATGAAGGTAGAGACACGCTCCATTAATAGTCCAG	1698

1458 TCTTTCAAGACCTGATGCTGCTGCTGCCGAACATGAAGACATGTCCAGTAATGAGTCGAG 1517

1699 AAATGCAACAGATGAATCCCTTCAGAAAGAGGTGATTGCAAAATCTGGCTGAGCATAATTCT 1758

1518 AAACGGCGGACAGACGAGCCCTTCGAGAGGAGGTGATCGCAAACTTGGCGGAGCACATTCT 1577

1759 ATTCACTGCTAGCAAGTCTGTGTGCCATTATGTCCACACACATTTGGTGTGCTGCTCCT 1818

1578 CTTTCAACGGCAGCAAGTCTGCGCTATCATGTCCACCCACATTTGTGGCTGCTGCTCCT 1637

1819 CTAACAGACACAGGCAAGGAATTGATCTCTCCACATTTGGTGAAGACTTCTTTTGTGATGAA 1878

1638 CTAACAGACACAGGCAAGGAATCCACTCTCCAGCTGTGGGAAGACTCTTTTGTGATGAA 1697

1879 AGAGGAAGTCTGGCTCGTGATTTTGACCTGGGGTTCTCAGAGAAATTCAGAAGATGTAGT 1938

1698 GGAGGAAGTCTTAGCTCGGATTTTGACCTGGGCTTCTCCGGAATTCAGAAGATGTAGT 1757

1939 AATGATGCCATACAGCTGCTGGGAAATTTGTGTCAATCACCCACACTAGCAGGAAACGA 1998

1758 CATGATGCTATTAGCTTCTGGGGAACGTGTGTCAATCACCCACACTAGCAGGAAAGGA 1817

1999 TGAGTTTTTATCACCCCCAGCAGACAACCTGTCCTCATCAGTCTTCGAACCTCAACTTCTACAG 2058

1818 TGAATTTCTTTATTACTCCAGCACAACCTGTCCGCTCCGTCTTTTGAACCTCAACTTCTACAG 1877

2059 CAACTGGGTACTTCATGCTTTTATCATGGAGGCCATCATAGCTTCGACGCTTATGCACT 2118

1878 CAACTGGGTACTTCATGCTTTTATCATGGAGGCCATCATAGCTTCGACGATTTATGCACT 1937

2119 TCTTGAACAGAGGGGACTTGGGGGTCCCACTAGCACCCCACTAACTGATCAGCCAGGA 2178

1938 CCAGAAATAGAGGGGTTCCGGAGGGTCTGCGGAGGCTTGGCAACCTGATCAGCCAGGA 1997

2179 GCAGCTGTGTGCGGAAGGCGGACGCTGTGCTACTCTTCTCTCAATGAAGGCAACCATCTC 2238

1998 GCAGCTGTGTGCGGAAGGCGGACGCTGTGCTACTCTTCTCTAAATGAAGGTACCATTTTC 2057

2239 ACTGCTTCCAGACATTTTACCAAGTCTGCCATGAACAGTAGAGAAAGTTTATCCAGTA 2298

2058 TCTGCTCCAGACATTTTACAGGTTTGTCAAGAGACAGTAGAGAAAGTTTATCCAGTA 2117

2299 TGGCATTTCTTACAGTGGCAGACAGCATGACACGAGGAAGATATCAGTCCCTAGTCTTGCTGA 2358

2118 CGGAATTTCTCAGTGGCAGAGCAAGATGACACGAGGAAGATGTCAGTCTCGGCTTGCAGA 2177

2359 GCACAGTGGGACAAGAGCTTCCAGAACCTTTGTCTTGGAGAAGTGATGAAGAGATGA 2418

2178 GCACAGTGGAAACAGAAAGCTTCCGAGACCTCTGAACTGGAGAAGTGACGAAGAAGATGA 2237

2419 AGACAGTGAATTTGGGGAGGACAGCGAGATGCTACTCTGAAGGTGAGCCCAATCCAGGA 2478

2238 GGAAGTGAATTTGGTGAAGAGACGCGTGAATGCTACTCTGAAGGTGAGCCAGGCGCAAGGA 2297

2479 GCACAGCAGATTTATCACTTTTACAGAGACTCTTGGGCTTTGCTGGAGGCTACAG 2538

2298 GCACCAACATTCATCACCTTTCTGACAGGCTTCTGGGGCCCCCTGCTTAGAAGCCTACAG 2357

2539 CTTGCTGCAATCTTTTGTTCACAACTTCAGTGGTCTGTTTCCAGAACCTGAGTATCTGCA 2598

2358 CTTGCTGCACTTTTGTGCCACACTTTCGCGGCCAGTCCCGGAGTCTGAGTACCTTGCAC 2417

2599 AAGTTGCACAAATACCTAATACCAGAACAGAAAGATTTGCAAGTATATGCTGAGAG 2658

2418 GAAGCTGCACAGGTACCTTCTCACAGGACGGAAGGAAACGTCGCGGGTGTACGCTGAGAG 2477

2659 TGCACATATGCTTTGTGAAGAAATGCTGTGAATAATGTTTAAAGGATATTTGGGGTTTTCAA 2718

2478 TGCACATACTGCTTTGTGAAGAAATGCTGTGAATAATGTTTAAAGGACATCGGGTTTTCAA 2537

2719 GGAGACCAACAAAGAGAGGTGCTGTTTTTGAACCTGAGACGACATTTTCTACTCTCAATG 2778

Db 2538 AGAGACCAAGCAGAGCGGCTGTCTTAGAACTGAGCACCACCTTCCTACCTCAGGG 2597  
Qy 2779 CAACCGCAAAACTCTAGATATATCTGAGTTTGTGGTGTCTAG 2827  
Db 2598 CAGCGGCGAGAGCTCCTGGAATACATCTGAGCTTCGTGTCTGTAG 2646

## RESULT 13

AF021348 2646 bp mRNA linear ROD 24-NOV-1999  
LOCUS Rattus norvegicus sn-glycerol-3-phosphate acyltransferase (GPAT)  
DEFINITION mRNA, nuclear gene encoding mitochondrial protein, complete cds.

ACCESSION AF021348  
VERSION AF021348.1 GI:2444458

KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus

1 (bases 1 to 2646)  
Ganesh Bhat, B., Wang, P., Kim, J. H., Black, T. M., Lewin, T. M.,  
Fiedorek, F. T. Jr. and Coleman, R. A.

TITLE Rat sn-glycerol-3-phosphate acyltransferase: molecular cloning and  
characterization of the cDNA and expressed protein  
JOURNAL Biochim. Biophys. Acta 1439 (3), 415-423 (1999)

MEDLINE 99376617  
PUBMED 1046428

REFERENCE 2 (bases 1 to 2646)  
Bhat, B. G., Wang, P., Kim, J. H., Black, T. M., Lewin, T. M.,  
Fiedorek, F. T. Jr. and Coleman, R. A.

AUTHORS Direct Submission  
TITLE Submitted (26-AUG-1997) Nutrition, Univ. of North Carolina, CB7400  
JOURNAL McGavaran-Greenberg Bldg., Chapel Hill, NC 27599, USA

FEATURES Location/Qualifiers  
1..2646

/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/tissue\_type="liver"

gene 1..2646  
/gene="GPAT"

CDS 160..2646  
/gene="GPAT"

/EC\_numbers="2.3.1.15"  
/note="similar to protein encoded by Mus musculus glycerol  
3-phosphate acyltransferase gene, GenBank Accession Number  
W77003, and Accession Numbers U36771, U36772, U36773"

/codon\_start=1  
/product="sn-glycerol-3-phosphate acyltransferase"  
/protein\_id="AA071605.1"

/db\_xref="GI:2444459"

/translation="MEESVITGIDVSYLNSSEYSIGRCCKHTNEDWDVCGPPTFP  
RSATLKWMSKRPVGRCCYCTPSWREFNPSISGLRNVIYINETHRR  
GWLARRSLYLFVQDERVHKGFMATISDNLNSRVQEAIAEVAALPDGSAQOOS  
KAIQKVRKARIQEMVATSPGMIRLTGWLLKLFNSFFWNIQIHKQLEWKAAT  
ETNLLPLFVPHRHIDYLLLFILFCHNIKAPYIAGNNLNIPIFSTLIHLGGLFFI  
RRLEDETPDGRKDILYLLHGHIVELLRQOFLIELEGTSSRSKTSACARGLLSV  
VDTLSSNTIPDILVPIVGSYDRIIEGHNGEQLKPKKNSLWSVARGVIRLKN  
GYRVDPAPQPSLUKEYLEQSPVAPLSLEQALLPAILPSRPDAADAEHEDMSN  
ESRNADEAFERRLIJANLAELHIFATSKSCAIMSTHIVACILLYRHRQGIHLSTLVED  
FVNMKEVLAERDFDIFGSGNEDVMHAIQLLCNVITHTSRKDEFITPTSTVPSV  
FELNFSYNGVLVHFIMEAIACSIYVONKRGSGSAGLGNLISQOLVRKASLCY  
LLSNEGTLISLPQFYQVQCKFYQYGLITVAEQDDQEDVSPGLAEQWKKLPE  
PLNRSDEEDSDFEQRQCYLKSQAKHQFIFLQRLGLLEAYSSAAIFVH  
TFRGVPESEYILKHLRYLLRTERNVAAYEATYCLVKNVRFKIDIGVFKETKQ  
RASVLELSLTFPQGRQKLELYLSFVL"

BASE COUNT 670 a 689 c 670 g 617 t  
ORIGIN

Query Match 63.1%; Score 1896.2; DB 10; Length 2646;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2126; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 319 TTTGGGAATTACACTTTTGTGACATGGATGAATCTGCACTGACCCCTTGGTGAACAATAGATGT 378  
Db 138 TTTGGGACTTCCAGTGTGAGTTCATGGAGGAGTCTTCTAGTGACAAATGGCAACAATAGAGT 197  
Qy 379 TTTCTATCTGCCACATTCATCAGAAATACAGTGTGTGGTGCATGTAAAGCAACAAGTGAGGA 438  
Db 198 TTTCTATCTGCCCAATTCATCAGAAATACAGCTTTGGCCGATGTAAACACACGAATGAGGA 257  
Qy 439 ATGGGGTGAAGTGTGGCTTTAGACCCACCGTCTTCAGATCTCGAACTTAAATGGAAGA 498  
Db 258 CTGGGTGACTGTGGCTTCAAAACCTTACCTCTTCAGATCCGCAACGCTGAAATGGAAGA 317  
Qy 499 AAGCCTAATGAGTCGGAAGGCCATTTGTTGGAGAGATGTTTACTCTCTGCACCTCCCA 558  
Db 318 GAGCTCATGAGCGGAAGAGGCCCTTCGTGGGAAGGTGTTGCTATTATCATGCACGCTCA 377  
Qy 559 GAGCTGGGACAAATTTTCAACCCAGTATCCCGTCTTTGGGGTTTGGCGAATGTTATTATA 618  
Db 378 GAGCTGGGAAAGGTTTTTCAACCCAGTATCCCATCTCTGGGTTTGGCGAATGTTATTATA 437  
Qy 619 TATCAATGAAACTCACACAGACACCGCGGATGCTTGCAGAGCGCTTTCTTACGTCT 678  
Db 438 TATCAATGAAACTCACACAGGACCGAGGATGCTTGCAGAGCGCTTTCTTACATCCT 497  
Qy 679 TTTTATTCAAGAGCGAGATGTGATTAAGGGCATGTTTGCACCAATGTGACTGAAATGT 738  
Db 498 TTTTGTTCAGAGCGCGATGTCCACAGGGCATGTTTGCACCAAGTATCATGACAATGT 557  
Qy 739 GCTGAACAGCAGTAGATGACAAAGGCAATTTGCAAGAGTGGCTGTGAATTAACCTTGA 798  
Db 558 ACTGAATAGCAGCAGAGTCCAAAGAGCAATTTGCTGAAGTGGCTGCAGAAATTTGAACCCGA 617  
Qy 799 TGGTCTCTCCCGCAGCAGCAATCAAAAGCGGTAAACAAGTGAAGAAAGAGCTAAAGAT 858  
Db 618 TGGATCTCTCCCGCAGCAGCATCAAAAGCCATCCAGAAAGTGAAGAAAGAGCAGGAAGAT 677  
Qy 859 TCTTCAAGAAATGTTGCCACTGTCTCACCAGCAATGATCAGACTGACTGGTGGTGGTCT 918  
Db 678 CTTCCAGGAATGTTGCTACAGTCTCCCCGGGATGATCAGGCTGACTGGCTGGGTGTT 737  
Qy 919 GCTAAACTGTTCAACAGCTTTCTTTTGGAAATTTCAAAATTCACAAAGGTCAACTTGAGAT 978  
Db 738 ACTAAAGCTCTTCAACAGCTTTCTTCTGAAATTCAGATTCAACAGGGTCAACTTGAGAT 797  
Qy 979 GGTAAAGCTGCAACTGAGCAGCAATTTGCCGCTTCTGTTTCTACAGTTCATAGATCCCA 1038  
Db 798 GGTAAAGCTGCAACTGAGCAGCAATTTGCCGCTTCTGTTTCTGCCGGTGCACAGATCCCA 857  
Qy 1039 TATTGACTATCTGCTGCTCACTTTTCTCTTCTGCCATAAATCAAGACCAATACAT 1098  
Db 858 CATGCACTACTGCTGCTCACTTTCTCTTCTGCCATAAATCAAGACCTCAATACAT 917  
Qy 1099 TGCTTCAGGCAATTAATCTCAACATCCCAATCTTCAGTACCTTTGATCCATAAGTTGGGG 1158  
Db 918 CGCTTCGGGCAACAACCTCAACATCCCATCTTTCAGTACCTTTGATTCACAAGTTGGGG 977  
Qy 1159 CTTCTTTCATCAGCAAGAGCTCGATGAACACACAGATGAGCGGAAGAGATTTCTCTATAG 1218  
Db 978 CTTTTCATAGACGAGGCTTGACGAATCTCCAGATGGACGCAAGACATTTCTGTACAG 1037  
Qy 1219 AGCTTTCCTCCTCAGGCAATAGTTGAATTTACTTTCGACAGCAGCAATTTCTTGGAGATCTT 1278  
Db 1038 AGCGTTCTCCTCAGGCAATAGTTGAATTTACTTTCGACAGCAGCAATTTCTTGGAGATCTT 1097  
Qy 1279 CCTGGAAGGACACGTTCTTAGGAGTGGAAAACCTCTTGTCTCGGGCAGGACTTTTGTCTC 1338  
Db 1098 CCTGGAAGGACACCGCTCCCGCAGTGGCAAGACCTCTGTCGCCCGGGCGGGCTCCTGTC 1157  
Qy 1339 AGTTGTGGTAGATACTCTGTCTTCAATGTCTATCCAGATCCCAAGACATTTGTATATACCTGTGG 1398  
Db 1158 AGTGTAGTGGATACTCTGTCTCATCAACACCATCCCTGACATCCTGTGTCATCCTCCTGTGGG 1217

QY 1399 AATCTCTATGATCGCATTTATGGAAGTCACTAAATGGTGGAACAACATGGGCAAACTAA 1458  
Db 1218 CATCTCGTATGATCGGATTAATCGAAGTCACTAACAATGGTGAACAGTTGGGCAAGCCAA 1277  
QY 1459 GAAGAATGAGAGCCTGTGGAGTGTAGCAAGAGTGTATTTAGAAATGTTAGCAAAAACATA 1518  
Db 1278 GAAGAATGAAAGTCTCTGGAGTGTGCAAGAGGCGTTATCAGAAATGCTGGGAAAAACTA 1337  
QY 1519 TGGTGTGTCGAGTGTGATTTTCACAGAGCCATTTCTTAAGGAATATTTAGAAAGCCA 1578  
Db 1338 CGGCTATGTCGAGTGTGATCTTTCGACAGCCATTTCTTGNAGAAATATTTAGAGGCCA 1397  
QY 1579 AAGTCAGAAACCGGTGTCTGCTCTACTTTCCCTGGAGCAAGCGTTGTTACAGCTATACT 1638  
Db 1398 AAGTCAGAAACCTGTATCTGCTCCCTCTCTTTGGAGCAAGCACTGTTACCAAGCAATCCT 1457  
QY 1639 TCCTTCAAGACCCAGTGTGCTGCTGATGAGGTAGAGACACGTCATTAATGAGTCCAG 1698  
Db 1458 TCCTTCAAGACCCAGTGTGCTGCTGCTGCGAACAATGAAGACATGTCAGTAATGAGTCCAG 1517  
QY 1699 AATGCAACAGATGAATCCCTACGAGGAGGTGATTGCAAAATCTGGCTGAGCATATTCT 1758  
Db 1518 AAACGGGCGACAGCAAGCCCTTCGAAAGGAGGCTGATCGCAAACTGGCGGAGCAATCT 1577  
QY 1759 ATTCACTGTCTAGCAAGTCTGTGCTGCTATGATGTCACACACATTTGTGGCTTGCCTGCTCT 1818  
Db 1578 CTTTACCGCCAGCAAGTCTGTGCTGCTATCATGTCCACCCACATTTGTGGCTGCTGCTCT 1637  
QY 1819 CTACAGACACAGCAGGGAATGATCTCTCCACATTTGCTGGAAGACTTTCTTGTGATGAA 1878  
Db 1638 CTACAGACACAGCAGGGAATCACCCTCTCACGCTGGTGGAAAGACTTTCTTGTGATGAA 1697  
QY 1879 AGAGGAAGTCTTGGCTGCTGATTTTGCACCTGGGGTCTCAGGAAATTCAGAAATCTGAT 1938  
Db 1698 GGAAGAGTCTTGTGCTGCGGATTTTGTGCTGGGCTTCTCGGAAATTCAGAAATCTGAT 1757  
QY 1939 AATGATGCCATACAGTGTGCTGGAATTTGTGTCACAAATCACCACACATAGCAGGAACGA 1998  
Db 1758 CATGATGCTATTTACGCTTCTGGGAACTGTGTACAAATCACCACATAGCAGGAAGGA 1817  
QY 1999 TGAGTTTTTTATCACCCGACGACAACTGTCCCATGAGTTCGAACTTCGAACTTCTACAG 2058  
Db 1818 TGAATTTCTTTATTAATCCAGCAAACTGTCCGCTCGCTTTTGAATCTCAACTTCTACAG 1877  
QY 2059 CAATGGGTACTTTCATGCTTTATCATGAGGCTCATCATGCTTGCAGCTTTATGCACT 2118  
Db 1878 CAATGGGTACTTTCATGCTTTTATCATGAGGCTCATCATGCTTGCAGCAATTTATGCACT 1937  
QY 2119 TCTGAAACAAGAGGGGACTGGGGGTCCCACTAGCACCCCACTTAACCTGATCAGCCAGGA 2178  
Db 1938 CCAGAAATAGAGGGGTTCCGAGGGTCTGCGGAGGCTTGGCAACTGATCAGCCAGGA 1997  
QY 2179 GCAAGTGTGTCGGAAGCGGCCAGCTGTGCTACTTCTTCTCAATGAAGGCAACATCTC 2238  
Db 1998 GCAAGTGTGTCGGAAGCGGCCAGCTGTGCTACTTCTTCTTAATGAAGGTACCACTTC 2057  
QY 2239 ACTGCTTGCACACATTTTACCAGTCTGCCATGAACAGTAGGAAAGTTTATCCAGTA 2298  
Db 2058 TCTGCTTGCACACATTTTACCAGGTTTGTCAAGAGACAGTAGGAAAGTTTATCCAGTA 2117  
QY 2299 TGGCAATTTTACAGTGGCAGACGATGACACGAGGATATCATGCTCTAGTCTTGTGTA 2358  
Db 2118 CGGAATTTCTACAGTGGCAGACGATGACACGAGGATGTCAGTCTTGGCTTGCAGA 2177  
QY 2359 GCAAGTGTGGAACAAGAGTTCAGAACCTTTGTCTTGTGAGAAAGTGTAGGAAGATGA 2418  
Db 2178 GCAAGTGTGGAACAAGAGTTCAGAACCTTCCGAGGCTCTGAACTGAGAAAGTGAACGAAGATGA 2237  
QY 2419 AGACAGTGTGTTGGGAGGAAACAGCGAGATTCCTACCTGAAGTGTGAGCAATCCAGGA 2478  
Db 2238 GGAAGTGTGTTGGTGTGAGGAGGAGGCTGATTCCTACCTGAAGTGTGAGCCAGGCAAGGA 2297  
QY 2479 GCACAGCAGTGTATCACCTTCTTACAGAGACTCTTGGGCTTTTGTGGAGGCTTACAG 2538

Db 2298 GCACCAAGCAATTCATCACCTTTCTGCAGAGGCTTCTGGGGCCCTCTAGAACCTACAG 2357  
QY 2539 CTCTGCTGCCATCTTTGTTTCAAACTTCAAGTCTGCTGCTTCCAGAACCTCTGAGTATCTGCA 2598  
Db 2358 CTCTGCTGCCATCTTTGTTTCAAACTTCCAGGCTTCCGCGCCAGTCCCGGAGTCTGAGTACCTGCA 2417  
QY 2599 AAAGTTGCAAAATACCTAATAACCAAGAACAGAAATGTTTGCAGTATATCTCTGAGAG 2658  
Db 2418 GAAGCTGCACAGGTACCTTCTCCAGGACGAGAGAACGTCGCGGTGTACGCTGAGAG 2477  
QY 2659 TGCACATATTTGTTGTGAAGATGCTGTGAAAAATGTTTAAAGATATTGGGGTTCCTCAA 2718  
Db 2478 TGCACATATTTGTTGTGAAGATGCTGTGAAAAATGTTTAAAGACATCGGGGTTTCCTCAA 2537  
QY 2719 GGAGCCAAACAAAGAGAGTCTGTTTAAAGCTAGACAGCAGCATTCTTACCTCAATG 2778  
Db 2538 AGAGCAACAGCAGAGCGAGCTGTCTTAGAACTGAGACCACTTCTTACCTCAGGG 2597  
QY 2779 CAACCCGACAAAACCTTCTAGAAATATATTCTGAGTTTGTGGTCTGTAG 2827  
Db 2598 CAGCCGGCAGAGCTCTCTGAAATACATTCTGAGCTTCTGTTGCTGTAG 2646

RESULT 14  
RNU36771  
LOCUS  
DEFINITION  
Rattus norvegicus sn-glycerol 3-phosphate acyltransferase mRNA, nuclear gene encoding mitochondrial protein, complete cds.  
ACCESSION  
U36771 U36773  
VERSION  
U36771.2 GI:4481754  
KEYWORDS  
Rattus norvegicus (Norway rat)  
SOURCE  
Rattus norvegicus  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
REFERENCE  
1 (bases 1 to 2519)  
AUTHORS  
Nikonov, A.V., Morimoto, T. and Haldar, D.  
TITLE  
Properties, purification and cloning of mitochondrial sn-glycerol 3-phosphate acyltransferase  
(in) Pandalai, S.G. (Ed.);  
JOURNAL  
RECENT RESEARCH DEVELOPMENTS IN LIPIDS RESEARCH, VOLUME 2: 207-222;  
Transworld Research Network, India (1998)  
REFERENCE  
2 (bases 1 to 2519)  
AUTHORS  
Nikonov, A.V., Morimoto, T. and Haldar, D.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (20-SEP-1995) Andrei V. Nikonov, Department of Biological Science, St. John's University, Jamaica, NY 11439, USA  
REFERENCE  
3 (bases 1 to 2519)  
AUTHORS  
Nikonov, A.V., Morimoto, T. and Haldar, D.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (23-MAR-1999) Cell Biology, New York University School of Medicine, 550 1st Ave. MSB-697, New York, NY 10016, USA  
REMARK  
Sequence update by submitter  
COMMENT  
On or before Mar 23, 1999 this sequence version replaced gi:1754786, gi:1754788, gi:1754784.  
FEATURES  
Location/Qualifiers  
1..2519  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague Dawley"  
/db\_xref="taxon:10116"  
/sex="male"  
/tissue\_type="liver"  
/clone\_lib="Stratage cat. number 936513"  
101..2419  
/EC\_number="2.3.1.15"  
/note="similar to product encoded by Mus musculus glycerol 3-phosphate acyltransferase gene, GenBank Accession Number M77003"  
/codon\_start=1  
/product="sn-glycerol 3-phosphate acyltransferase"



```

/protein_id="AAB39470.2"
/db_xref="GI:4481755"
/translation="MSRRKPFVGRCCYSCSTPQSWERFPNFPPIPLGLRNVIYINETH
RQGWLARLSYILFVQRDVHKGWFATSIIDNVLNRSVQEAIAEVAALNIPDQSAV
QSKAIQKWKARKILQEMVATVSPGMRILTQVLLKLFNSFFWNIQIHKGOLEMYK
AATNPLPLPVRSHIDYLLATFILLFCHNIKAPYIASGNLNIPIFSLIHLKGG
FFIRRLDTPDGRDILYRALLGHVVVELLQQOFLIFLEGFRSRGKTSFCARAGV
LSVVDTLSSNTIPDILIPVIGISYDRIIEGHYNGEQLGPKKNESLWSVARGVIRML
KNYGVYRVDAQFSLKEYLEGOSQKPSVAPLSLEQALLPAIILPSRPDAAAHEHDM
SSNRSNAADFRRRLITANLAENHLLTASCSAIMTHIVACILLYRHWOGLHSLTL
VEDEFVKEEVLNDFDLFGSNGSEDDVMHAIQILGNCTVITHTSRKDERPITPSTTV
PSVPELNFYSNGVDFVMEALIASIVAVONKSGSLPEPGLNLSQEQQLYRAASL
CYLLSNEGTISLPQTFQVQCHETVGFQIIGILTVAEQDDQEDVSFGLAEQWNNKL
PEPLNWSDEDESDSDFGEQRDCYLKVSQAKHQQFITFLQRLVGLLLEAYSSAAIF
HTFAPVPPESEYIQLKRLHYLLTRTERNAVAYAESATYCLVKNVAKMFKDITGVFKETK
QKRASVLELSYTFLLPQSRQKLLVILSFVL"
transit_peptide 101..286
misc_feature 1367..1405
/evidence=not_experimental
misc_feature 1658..1705
/evidence=not_experimental
misc_feature 254..286
/evidence=not_experimental
/note="encodes R(-10) motif recognizable by matrix
proteases"
BASE COUNT 627 a 670 c 644 g 578 t
ORIGIN
Query Match 60.2%; Score 1807.2; DB 10; Length 2519;
Best Local Similarity 84.2%; Pred. NO. 0;
Matches 2050; Conservative 0; Mismatches 383; Indels 3; Gaps 1;
QY 406 CAGTGTTCGATGATGACACACAAAGTGGGATGGGTGAGTGGCTTTAGACCCAC 465
DB 1 CAGCGTTGGCGATGTAACACACAGATGAGGACTGGGTGACCTTTGGCTTCAACCTAC 60
QY 466 CCGTCTTCAGATCTGCAACTTTAAATGAAAGAAAGCCCTAATGATCGGAAAGGCCATT 525
DB 61 CTTCTTCAGATCCGCNACGCTGAAATGGAAGAGAGCCCTCATGAGCGGAGAGGCCCTT 120
QY 526 TGTGGAAAGATGTTGTTACTCTCGACCTCCCGAGCTGGGACAAATTTTCAACCCCGAG 585
DB 121 CGTGGGAAGTGTGCTATTATCATGACGCTCAGAGCTGGGAAAGGTTCTTCAACCCCGAG 180
QY 586 TATCCCGTCTTGGGTTTGGGATGTTATTTATATCAATGAAACTCACAGACACCG 645
DB 181 TATCCCGCTCTGGGTTTGGGAAATGTTATTTATATCAATGAAACTCACAGACACCG 240
QY 646 CGGATGGCTTCAAGACGCCCTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTGCATAA 705
DB 241 AGGATGGCTGCAAGACGGCTTTCTTACATCTTTTGTTCAGAGCGCGATGTCCACAA 300
QY 706 GGGCATGTTTCCCAACCAATGTACTGAAAATGTGCTGAACAGCAGTACAGAGGCG 765
DB 301 GGGCATGTTTCCCAACCAATGTACTGAAAATGTGCTGAACAGCAGTACAGAGGCG 360
QY 766 AATTGCAGAGTGGCTGCTGAATTAACCCGTATGTTTCTGCCAGCAGCAGCAATCAAAAGC 825
DB 361 AATTGCTGAAGTGGCTGAGAAATTAACCCCGATGGATCTGCCAGCAGCAGTCCAAAGC 420
QY 826 CGTTAAACAAAGTGAAGAAAGAGCTTAAAGATTTCTTCAAGAAATGGTTGCCACTGCTC 885
DB 421 CATCAGAAAGTGAAGAAAGCAGGAGATCTCCAGGAAATGGTTGTCAGTCTC 480
QY 886 ACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTAAACCTGTTTCAACAGCTTCTTTTG 945
DB 481 CCCCAGGATGATCAGGCTGACTGGCTGGGTGTTACTAAAGCTCTTCAACAGCTTCTCTG 540
QY 946 GAACATTCAAATTCACAAAGTCAACTTGAGATGGTTAAAGCTCAACTGAGACGAATTT 1005
DB 541 GAACATTCAGATTCACAAAGGTCAACTTGAGATGGTGAAGCTGCAACTGAGACGAATCT 600

```

```

QY 1006 GCCGCTTCTGTTTCTTACCAGTTTATAGATCCCATATTAAGTACTATCTGTGCTCACTTTTCAT 1065
DB 601 GCCGCTCTGTTTCTGCGGGTGACAGATCCCATCGACTACCTGTTGCTCACTTTCAT 660
QY 1066 TCTCTTCTGCCATTAACATCAAGCACCATACATTCGTTTCAGGCAATTAATCTCAACATCCC 1125
DB 661 CCTCTTCTGCCAACAATCAAGCTCCATACATCGCTCGGCAACCACTCTCAACATCCC 720
QY 1126 AATCTTCAGTACCTTGTATCATAAGCTTGGGGGCTTCTTATACAGCAGGAGCTCGATGA 1185
DB 721 CATCTTCAGTACCTTGTATCATAAGCTTGGGGGCTTCTTATACAGCAGGAGCTTGA 780
QY 1186 AACACAGATGGAACGGAAGATGTTCTTATAGAGCTTTCCTCATGGGCAATAGATTGA 1245
DB 781 AACTCCAGATGGACGCAAGACATTTCTGTACAGAGCGTTGCTCCATGGGCATGTATTGA 840
QY 1246 ATTACTTCGACAGCAGCAATTTCTGGAGATCTTCTGGAAGGACACGTTCTAGAGTGG 1305
DB 841 ACTCTCCGACAGCAGCAGTTCCTGGAGATCTTCTGGAAGGACCCGCTCCCGCAGTGG 900
QY 1306 AAAAACCTCTTGTCTCGGCGAGCCTTTTGTCAAGTTGTGTAGATATCTCTGTCTACCAA 1365
DB 901 CAGACCTCTGTGCTCCCGGCGCGCTCTGTCAAGTGTGTAGTGTATCTCTGTCTATCCAA 960
QY 1366 TGTCAATCCGACATCTTGTATATACTGTGTGAATCTCTTATGATCGCATTTATCGAAGG 1425
DB 961 CACCATCCCTGACATCTGTCATCTCTGTGGGATCTCTGTATGATCGGTAATCGAAGG 1020
QY 1426 TCATCAATAGTGTGAACAACTGGGCAAACTTAAGAAATGAGAGCCTGTGGAGTGTAGC 1485
DB 1021 TCACTACAATAGTGTGAACAACTGGGCAAGCCCAAGAAATGAAAAGTCTCTGAGTGTGGC 1080
QY 1486 AAGAGGTGTATTAGATGTTACGAAAACCTAAGTGTGTGTCTCGAGTGGATTTTGACA 1545
DB 1081 AAGAGGCTTATCAGAATGTGCGGAAAACTACGGCTATGTCTCGAGTGGACTTTTGACA 1140
QY 1546 GCCATTTTCTTAAAGAAATATTATTAAGAACCAAGTCAAGAACGGTGTCTGTCTACT 1605
DB 1141 GCCATTTTCTTGAAGAAATATTATTAAGGCCAAGTCAAGAACTGTATGTCTCCCT 1200
QY 1606 TTCCCTCGAGCAAGCGTTGTATCCAGCTATATCTTCTTCAAGACCCAGTGTGTCTGA 1665
DB 1201 CTCTTTGGAGCAAGCAGTGTACAGCAATCTTCTTCTTCAAGACCTGTATGTCTGTCTGC 1260
QY 1666 TGAAGTGAACACACGCTCCATTAATGATCCAGAAATGCAACAGATGAATCCCTACGAG 1725
DB 1261 CGAACATGAAGACATGTCCAGTAAATGATCGAGAAACCGCGCAGACGAAAGCTTTCCGAG 1320
QY 1726 GAGGTTGATTTGCAATCTGCTGAGCATATTCTTATTCATCTGTAGCAAGTCTCTGCGCAT 1785
DB 1321 GAGGCTGATGCAAACTGCGGAGGACATCTCTCTACCGCCAGCAAGTCTCTGCGCTAT 1380
QY 1786 TATGTCCACACACATTTGCTGCTGCTCTCTACAGACACAGCAGGAGGAAATGATCT 1845
DB 1381 CATGTCCACCCACATTTGTCGCTGCTCTCTACAGACACCTGCGCAGGAGAAATCCACCT 1440
QY 1846 CTCCACATTTGGTTCGAAAGCTTTTGTGTAGTGAAGAGAGAGTCTCTGGCTCGTGAATTTGA 1905
DB 1441 CTCCACGCTGGTGGAGACTTTCTTGTGTAGTGAAGAGAGTCTCTAGCTCGGGAATTTGA 1500
QY 1906 CTTGGGCTTCTCAGGAAATTCAGAAAGATGTAGTAATCATGCCCATACAGCTGCTGGGAAA 1965
DB 1501 CTTGGGCTTCTCAGGAAATTCAGAAAGATGTAGTAATCATGCCCATACAGCTTCTGGGCAA 1560
QY 1966 TTGTGTCACAATCAACCACTAGCAGGAAAGTGTATTTTATCAACCCCGCAGCAAC 2025
DB 1561 CTGTGTCACAATCAACCCACACTAGCAGGAGGATGATTTCTTTATTACTCCACGACACAC 1620
QY 2026 TGTCCCATCAGTCTTCAAGCTCAACTCTTCAAGCAATGGGGTACTTCAATGTCTTTATCAT 2085
DB 1621 TGTCCCGCTCGTCTTTGAACTCAACTCTTCAAGCAATGGGGTACTTCAATGTCTTTATCAT 1680

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 18:05:10 ; Search time 515.003 Seconds  
(without alignments)  
15740.506 Million cell updates/sec

Title: US-09-935-290-1  
Perfect score: 3003  
Sequence: 1 ttccgaccaggctgctgcg.....ttcgagatctccagctttt 3003

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N\_Geneseq 13Jun03.\*
- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
  - 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
  - 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
  - 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
  - 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
  - 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
  - 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
  - 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
  - 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
  - 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
  - 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
  - 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
  - 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
  - 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
  - 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
  - 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
  - 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
  - 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
  - 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
  - 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
  - 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
  - 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
  - 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
  - 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*
  - 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3001	99.9	3003	24	ABK11094
2	2970.2	98.9	4031	23	ABV25313
3	2958	98.5	3210	24	AAJ35221
4	2591.4	86.3	2684	25	ABX72233
5	2485.4	82.8	2487	24	ABK94821
6	2483.8	82.7	2487	25	AAK55475
7	2334	77.7	3273	24	ABZ12025
8	1896.2	63.1	2646	24	ABK63741

9	437	14.6	469	22	ABA08415	Human sn-glycerol-
10	342.2	11.4	429	25	ABX42219	Bovine EST associa
11	337.6	11.2	407	25	ABX43656	Bovine EST associa
c	302	10.1	418	25	ABX35577	Bovine EST associa
12	246.4	8.2	271	21	AAC05673	Human secreted pro
13	246.4	8.2	387	25	ABX42469	Bovine EST associa
14	244.8	8.2	303	25	ABX47921	Bovine EST associa
c	233	7.8	264	25	ABX41074	Bovine EST associa
15	209.6	7.0	2000	24	ABK94822	Human glycerol-3-p
17	203.4	6.8	2000	24	ABK94822	Bovine glycerol-3-p
c	171	5.7	501	25	ABX38896	Bovine EST associa
19	161.8	5.4	231	24	ABN22191	Human ORFX polynuc
20	109	3.6	2682	24	ABN22191	cDNA encoding huma
21	109	3.6	2755	24	ABN22191	Human TRNFR-7 cDNA
22	103.8	3.5	3439	23	ABL09555	Drosophila melanog
c	103.8	3.5	9676	23	ABL09555	Drosophila melanog
23	93.6	3.1	399	25	ABX34889	Bovine EST associa
c	93.6	3.1	404	25	ABX40591	Bovine EST associa
25	83.2	2.8	404	25	ABX39377	Bovine EST associa
c	56.4	1.9	344	25	ABX39377	Bovine EST associa
27	53.8	1.8	2640	23	ABN36720	DNA encoding novel
28	49.2	1.6	60	24	ABN36720	Human spliced tran
29	49.2	1.6	65	24	ABN28887	Rat spliced trans
30	47	1.6	2502	22	AAO23332	E. coli plsb DNA e
31	47	1.6	2502	22	AAO23332	Escherichia coli p
32	46.2	1.5	2154	23	AAO23332	DNA encoding novel
33	46.2	1.5	2154	23	AAO23332	DNA encoding novel
34	46.2	1.5	2154	23	AAO23332	DNA encoding novel
35	45.6	1.5	2175	23	ABL07919	Drosophila melanog
36	45	1.5	100848	22	AAE28552	Genomic fragment #
c	44.8	1.5	1830121	17	AAE28552	Haemophilus influe
38	43.6	1.5	1127	23	AAE28552	DNA encoding novel
39	43.6	1.5	1127	23	AAE28552	DNA encoding novel
c	43.6	1.5	4120	23	AAE28552	DNA encoding novel
41	41.2	1.4	185371	24	ABT10718	Human breast cance
42	39.8	1.3	21045	22	AAE26721	Human genomic DNA
43	39.8	1.3	21045	25	ABX74070	Human novel polynu
c	39	1.3	1673	23	AAE26721	DNA encoding novel
c	39	1.3	2114	21	AAE26721	Human secreted pro

ALIGNMENTS

RESULT 1

ABK11094  
ID ABK11094 standard; cDNA; 3003 BP.

AC ABK11094;

XX 18-JUN-2002 (first entry)

DT cDNA encoding human acyltransferase, ACTR-1.

DE Human; acyltransferase, ACTR-1; antidiabetic; anorectic;  
KW Human; acyltransferase, ACTR-1; antidiabetic; anorectic;  
KW antiarteriosclerotic; cerebroprotective; anorectic; ophthalmological;  
KW cardiatic; metabolic disorder; energy homeostasis disorder; diabetes;  
KW hyperglycaemia; hypercholesterolaemia; hyperlipoproteinaemia; stroke;  
KW hypertriglyceridaemia; hyperlipidaemia; atherosclerosis; obesity;  
KW reitropathy; nephropathy; peripheral neuropathy; weight disorder;  
KW appetite regulation disorder; cachexia; anorexia; bulimia;  
KW cardiovascular disorder; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 341..2827  
FT FT /\*tag= a  
FT FT /product= "Human acyltransferase ACTR-1"  
FT FT /note= "The coding region is specifically claimed  
in claim 1"

XX WO200216592-A2.

XX 28-FEB-2002.



Db 1261 GCAATCTCTGGAGATCTTCTGGAAGGCACACAGTCTTAGAGTGGAAAAAAGCTCTGTGTC 1320  
Qy 1321 TCGGCGAGGACTTTTGTGTCAGTTGGTAGATACTCTGTCTACCAATGTCTATCCAGACAT 1380  
Db 1321 TCGGCGAGGACTTTTGTGTCAGTTGGTAGATACTCTGTCTACCAATGTCTATCCAGACAT 1380  
Qy 1381 CTTGATATACCTGTTGGAATCTCTATGATGCGCATTAATCGAAGTCTACTCAATGGTGA 1440  
Db 1381 CTTGATATACCTGTTGGAATCTCTATGATGCGCATTAATCGAAGTCTACTCAATGGTGA 1440  
Qy 1441 ACAACTGGGCAAACTTAAGAAAGATGAGAGCTGTGGAGTGTAGCAAGAGGTGTTATTAG 1500  
Db 1441 ACAACTGGGCAAACTTAAGAAAGATGAGAGCTGTGGAGTGTAGCAAGAGGTGTTATTAG 1500  
Qy 1501 AATGTTACGAAAAAATCTATGTTGTGTCGAGTGGATTTGACAGGCAATTTTCTTAAA 1560  
Db 1501 AATGTTACGAAAAAATCTATGTTGTGTCGAGTGGATTTGACAGGCAATTTTCTTAAA 1560  
Qy 1561 GGAATATTTAGAAAGCAAGTCTAGAAACCGGTGTCTGCTCTACTTTCCCTGGAGCAAGC 1620  
Db 1561 GGAATATTTAGAAAGCAAGTCTAGAAACCGGTGTCTGCTCTACTTTCCCTGGAGCAAGC 1620  
Qy 1621 GTTGTATACGCTATACCTTCTTCAAGACCCAGTGATGCTGCTGATGAAGGTAGAGACAC 1680  
Db 1621 GTTGTATACGCTATACCTTCTTCAAGACCCAGTGATGCTGCTGATGAAGGTAGAGACAC 1680  
Qy 1681 GTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTAGCAAGAGGTGATTTGCAAA 1740  
Db 1681 GTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTAGCAAGAGGTGATTTGCAAA 1740  
Qy 1741 TCTGCTGAGCATATCTTATCTACGTAGCAAGTCTGTGCGCATTAATGTCCACACACAT 1800  
Db 1741 TCTGCTGAGCATATCTTATCTACGTAGCAAGTCTGTGCGCATTAATGTCCACACACAT 1800  
Qy 1801 TGTGCTTGTGCTGCTCTACAGACAGCAGGGAATGATCTCTCCACATTTGTCGA 1860  
Db 1801 TGTGCTTGTGCTGCTCTACAGACAGCAGGGAATGATCTCTCCACATTTGTCGA 1860  
Qy 1861 AGACTTCTTTGTGATGAAGAGGAAGTCTCTGCTGCTGATTTGACCTGGGGTCTCAGG 1920  
Db 1861 AGACTTCTTTGTGATGAAGAGGAAGTCTCTGCTGCTGATTTGACCTGGGGTCTCAGG 1920  
Qy 1921 AAATTCAGAAGATGATGATGATGCAATGCAATGCAATGCTGCGGAAATTTGTGCAAAATCAC 1980  
Db 1921 AAATTCAGAAGATGATGATGCAATGCAATGCAATGCTGCGGAAATTTGTGCAAAATCAC 1980  
Qy 1981 CCACACTAGCAGGACGATGATGATTTTATCACCCCGAGCAACAGTGTCCCATCAGTCTT 2040  
Db 1981 CCACACTAGCAGGACGATGATGATTTTATCACCCCGAGCAACAGTGTCCCATCAGTCTT 2040  
Qy 2041 CGAACTCAACTTCTACAGCAATGGGGTACTTCTATGTCCTTTATCATGGAGGCCATCATAGC 2100  
Db 2041 CGAACTCAACTTCTACAGCAATGGGGTACTTCTATGTCCTTTATCATGGAGGCCATCATAGC 2100  
Qy 2101 TTGAGCCTTTATGCAAGTCTTGAACAGAGGGGACTGGGGGTCCCACTAGCAGCCCAACC 2160  
Db 2101 TTGAGCCTTTATGCAAGTCTTGAACAGAGGGGACTGGGGGTCCCACTAGCAGCCCAACC 2160  
Qy 2161 TAACTGATGAGCCAGGAGAGTGTGCGGAAGCGGCGGCGCTGTGCTACCTTCTCTC 2220  
Db 2161 TAACTGATGAGCCAGGAGAGTGTGCGGAAGCGGCGGCGCTGTGCTACCTTCTCTC 2220  
Qy 2221 CAATGAAGGCACCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCGCATGAACAGT 2280  
Db 2221 CAATGAAGGCACCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCGCATGAACAGT 2280  
Qy 2281 AGGAAAGTTTATCCAGTATGGCAATCTTACAGTGGCAGAGCAGATGACCAAGGAAGATAT 2340  
Db 2281 AGGAAAGTTTATCCAGTATGGCAATCTTACAGTGGCAGAGCAGATGACCAAGGAAGATAT 2340  
Qy 2341 CAGTCTTATGCTGTGAGCAGAGTGGGAACAAGAGCTTCCAGAACTTTGTCTTGGAG 2400  
Db 2341 CAGTCTTATGCTGTGAGCAGAGTGGGAACAAGAGCTTCCAGAACTTTGTCTTGGAG 2400

Qy 2401 AAGTGATGAAGAAGATGAAGACAGTACTTTTGGGGAGGAAACAGGAGATTTCTTACCTGAA 2460  
Db 2401 AAGTGATGAAGAAGATGAAGACAGTACTTTTGGGGAGGAAACAGGAGATTTCTTACCTGAA 2460  
Qy 2461 GGTGAGCAATCCAAAGAGAGCAGAGATTTATCACCTTTTACAGAGACTCTCTTGGGCC 2520  
Db 2461 GGTGAGCAATCCAAAGAGAGCAGAGATTTATCACCTTTTACAGAGACTCTCTTGGGCC 2520  
Qy 2521 TTTCTGAGGCTTACAGCTCTGCTGCATCTTTTGTTCACAACTTTTCTAGTGTCTGTCTCC 2580  
Db 2521 TTTCTGAGGCTTACAGCTCTGCTGCATCTTTTGTTCACAACTTTTCTAGTGTCTGTCTCC 2580  
Qy 2581 AGAACTCTGAGTATCTGCAAAAGTTGCAAAATCTTAATAACCAAGAACAGAAAGATGT 2640  
Db 2581 AGAACTCTGAGTATCTGCAAAAGTTGCAAAATCTTAATAACCAAGAACAGAAAGATGT 2640  
Qy 2641 TGCAGTATATGCTGAGAGTGCACATATTTGTGTGGAAGATGCTGTGAAATGTTTAA 2700  
Db 2641 TGCAGTATATGCTGAGAGTGCACATATTTGTGTGGAAGATGCTGTGAAATGTTTAA 2700  
Qy 2701 GGATATTTGGGCTTTTCAAGGAGACCAAAACAAAGAGAGTGTCTGTTTGTAGAACGAGCAG 2760  
Db 2701 GGATATTTGGGCTTTTCAAGGAGACCAAAACAAAGAGAGTGTCTGTTTGTAGAACGAGCAG 2760  
Qy 2761 CACTTTTCTACTCTCAATGCAACCGACCAAACTTTCTAGAAATATATTTCTGAGTTTGTGGT 2820  
Db 2761 CACTTTTCTACTCTCAATGCAACCGACCAAACTTTCTAGAAATATATTTCTGAGTTTGTGGT 2820  
Qy 2821 GCTGTAGGTAACGCTGGCCTGCTGCGCAATGAAGTCTATGATGAGTCTCTTGTAGG 2880  
Db 2821 GCTGTAGGTAACGCTGGCCTGCTGCGCAATGAAGTCTATGATGAGTCTCTTGTAGG 2880  
Qy 2881 TACCAGCTTCTGGCTCAAGAGTTTGAAGGTGCTTCCAGGGGTGAGGCTGCGCTGCTGTC 2940  
Db 2881 TACCAGCTTCTGGCTCAAGAGTTTGAAGGTGCTTCCAGGGGTGAGGCTGCGCTGCTGTC 2940  
Qy 2941 CGAAGTATCTCTCGAAGCAACAGTGTCTTCTNCTCCATGGATCTGAGATCTTCCAGC 3000  
Db 2941 CGAAGTATCTCTCGAAGCAACAGTGTCTTCTNCTCCATGGATCTGAGATCTTCCAGC 3000  
Qy 3001 TTT 3003  
Db 3001 TTT 3003  
RESULT 2  
ABV25313  
ID ABV25313 standard; cdna; 4031 BP.  
XX AC ABV25313;  
XX AC ABV25313;  
DT 16-SEP-2002 (first entry)  
XX Human prostate expression marker cdna 25304.  
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
OS Homo sapiens.  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.





QY 1683 CCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAAGGAGTTGAAATC 1742  
Db 1683 CCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAAGGAGTTGAAATC 1742  
QY 1743 TGGCTGAGCATATTCATTCTAGCTAGCAAGTCTGTGCCATTATGTCCACACACATTG 1802  
Db 1743 TGGCTGAGCATATTCATTCTAGCTAGCAAGTCTGTGCCATTATGTCCACACACATTG 1802  
QY 1803 TGGCTTGCTCTCTCTACAGACACAGCAGGGAATGATCTCTCCACATTGGTCCGAAG 1862  
Db 1803 TGGCTTGCTCTCTCTACAGACACAGCAGGGAATGATCTCTCCACATTGGTCCGAAG 1862  
QY 1863 ACTTCTTTGTGATGAAGAGGAGTCTTGGCTCGTGATTTTGACCTGGGGTTCTCAGGAA 1922  
Db 1863 ACTTCTTTGTGATGAAGAGGAGTCTTGGCTCGTGATTTTGACCTGGGGTTCTCAGGAA 1922  
QY 1923 ATTCAAGATGCTAGTAATGATGCATACATACAGCTGCTGGGAATTTGTCTCAATCACCC 1982  
Db 1923 ATTCAAGATGCTAGTAATGATGCATACATACAGCTGCTGGGAATTTGTCTCAATCACCC 1982  
QY 1983 ACATAGCAGGAACCATGAGTTTTTTATCACCCCCAGCACAACTGTCCCATCAGTCTTCG 2042  
Db 1983 ACATAGCAGGAACCATGAGTTTTTTATCACCCCCAGCACAACTGTCCCATCAGTCTTCG 2042  
QY 2043 AACTCAATCTTACAGCAATGGGTACTTCTATGTCTTTATCATGAGGCCATCATAGCTT 2102  
Db 2043 AACTCAATCTTACAGCAATGGGTACTTCTATGTCTTTATCATGAGGCCATCATAGCTT 2102  
QY 2103 GCAGCTTTATGCACTTCTGACAGAGGGGACTGGGGGTCCTCCACTAGCACCCACCTA 2162  
Db 2103 GCAGCTTTATGCACTTCTGACAGAGGGGACTGGGGGTCCTCCACTAGCACCCACCTA 2162  
QY 2163 ACCTGATCAGCAGGAGCAGTGTGCGGAAGGGCCAGCTGTGCTACCTTCTCTCCA 2222  
Db 2163 ACCTGATCAGCAGGAGCAGTGTGCGGAAGGGCCAGCTGTGCTACCTTCTCTCCA 2222  
QY 2223 ATGAAGGACCATCTCACTGCTTCCAGACATTTTACCAAGTCTGCCATGAACAGTAG 2282  
Db 2223 ATGAAGGACCATCTCACTGCTTCCAGACATTTTACCAAGTCTGCCATGAACAGTAG 2282  
QY 2283 GAAAGTTTATCCAGTATGGCATTCTTACAGTGGCAGACACATGACACAGAAATATCA 2342  
Db 2283 GAAAGTTTATCCAGTATGGCATTCTTACAGTGGCAGACACATGACACAGAAATATCA 2342  
QY 2343 GTCTAGTCTTGTCTGAGCAGTGGGACAAAGAGCTTCCAGAACTTTGCTTGGAGAA 2402  
Db 2343 GTCTAGTCTTGTCTGAGCAGTGGGACAAAGAGCTTCCAGAACTTTGCTTGGAGAA 2402  
QY 2403 GTGATGAAGAAGATGAAGACAGTGACTTTGGGGAGGAAACAGCGAGATTGCTACCTGAAGG 2462  
Db 2403 GTGATGAAGAAGATGAAGACAGTGACTTTGGGGAGGAAACAGCGAGATTGCTACCTGAAGG 2462  
QY 2463 TGAGCCTATCCAGAGACACAGCGAGTTTATCACTTCTTACAGAGATCTCTTGGGCTT 2522  
Db 2463 TGAGCCTATCCAGAGACACAGCGAGTTTATCACTTCTTACAGAGATCTCTTGGGCTT 2522  
QY 2523 TGCTGGAGCCTACAGCTCTGCTGCTCATCTTGTTCACAACTCAGTGGTCTCTGTCAG 2582  
Db 2523 TGCTGGAGCCTACAGCTCTGCTGCTCATCTTGTTCACAACTCAGTGGTCTCTGTCAG 2582  
QY 2583 AACCTGAGTATCTGCAAAAGTTGCAAAATACCTAATACCAAGAACAGAAAGAAATGTTG 2642  
Db 2583 AACCTGAGTATCTGCAAAAGTTGCAAAATACCTAATACCAAGAACAGAAAGAAATGTTG 2642  
QY 2643 CAGTATATGCTGAGAGTGGCACAATTTCTCTTGTGAAGAATGCTGTGAAAATGTTAAGG 2702  
Db 2643 CAGTATATGCTGAGAGTGGCACAATTTCTCTTGTGAAGAATGCTGTGAAAATGTTAAGG 2702  
QY 2703 ATATTGGGGTTTCAAGAGACCAACAAAGAGAGTGTCTGTTTAACTAGACAGCA 2762  
Db 2703 ATATTGGGGTTTCAAGAGACCAACAAAGAGAGTGTCTGTTTAACTAGACAGCA 2762

QY 2763 CTTTCTACCTCAATGCAACCGACAAAACCTTCTAGAAATATATTCTGAGTTTGTGTGC 2822  
Db 2763 CTTTCTACCTCAATGCAACCGACAAAACCTTCTAGAAATATATTCTGAGTTTGTGTGC 2822  
QY 2823 TGTAGGTAAACGTGTGGCACTGCTGGCAATGAAGTCAATGAGATCAGTTCCTTGTAGTA 2882  
Db 2823 TGTAGGTAAACGTGTGGCACTGCTGGCAATGAAGTCAATGAGATCAGTTCCTTGTAGTA 2882  
QY 2883 CCAGCTTCTGGCTCAAGAGTTTGAAGTGTCTTCGAGGGTCCAGGCTGCCCTGTCG 2942  
Db 2883 CCAGCTTCTGGCTCAAGAG-TTGAAGTGTCCGTCGCA-GGTCAGGCCCTGCCCTGTCG 2940  
QY 2943 AAGTGATCTCTGAAAGACAAAGTGCCTTCTNCCCTCCATGGAATCTGAGATCTTCCAGCTT 3002  
Db 2941 AAGTGATCTCTGAAAGACAAAGTGCCTTCTCCCTCCATGGATCTGTGATCTTCCAGCTC 3000  
QY 3003 T 3003  
Db 3001 T 3001  
RESULT 3  
AAD35221  
ID AAD35221 standard; cDNA; 3210 BP.  
XX  
AC AAD35221;  
XX AC  
XX 25-JUL-2002 (first entry)  
XX  
DE Human TRNFR-6 cDNA.  
XX  
KW Human; transferase; developmental disorder; cell proliferative disorder;  
KW TRNFR-6; neurological disorder; autoimmune disorder; parasitic infection;  
KW inflammatory disorder; endocrine; antiparasitic; immunosuppressive;  
KW cytostatic; neurological; gene; ss.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 338..2824  
FT /\*tag= a  
FT /product= "Human TRNFR-6 protein"  
XX  
FN WO200226950-A2.  
XX  
PD 04-APR-2002.  
XX  
PF 28-SEP-2001; 2001WO-US30424.  
XX  
PR 29-SEP-2000; 2000US-236523P.  
PR 06-OCT-2000; 2000US-238481P.  
PR 27-OCT-2000; 2000US-244025P.  
PR 03-NOV-2000; 2000US-246001P.  
PR 09-NOV-2000; 2000US-247931P.  
PR 16-NOV-2000; 2000US-249639P.  
PR 21-NOV-2000; 2000US-252819P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
XX Lal PG, Tang YT, Yue H, Burford N, Gandhi AR, Warren BA, Yao MG;  
PI Tribouley CM, Baughn MR, Lee EA, Hatalia AJA, Lu Y, Griffin JA;  
PI Sanjanwala MS, Ding L;  
XX  
DR WPI; 2002-362492/39.  
DR P-PSDB; AAE22144.  
XX  
XX Novel human transferase polypeptides and polynucleotides, useful in  
PT treating e.g., cell proliferative and autoimmune disorders -  
XX  
XX Claim 81; Page 154-155; 168pp; English.  
XX  
XX The present invention relates to novel human transferases (TRNFR) and  
CC polynucleotides encoding such proteins. The TRNFR proteins are useful  
CC

CC for treating disorders associated with a decreased expression of  
CC functional TRNFR, e.g., cell proliferative, developmental, neurological,  
CC autoimmune/inflammatory disorders and parasitic infections. Antagonists  
CC of TRNFR proteins are useful in treating disorders associated with  
CC increased activity of TRNFR. The present sequence is human TRNFR-6 cDNA.  
XX  
SQ Sequence 3210 BP: 896 A; 721 C; 735 G; 858 T; 0 other;

Db 1989 GGAACGATGAGTGTCTTTATACACCCCGACGAACTGTCCATCATCTTTCGAACTCAACT 2048  
Qy 2052 TCTACAGCAATGGGTACTTTCATGTCTTTATCATGAGGCCATCATAGCTTGCAGCCTTT 2111  
Db 2049 TCTACAGCAATGGGTACTTTCATGTCTTTATCATGAGGCCATCATAGCTTGCAGCCTTT 2108  
Qy 2112 ATGCAGTCTGAACAAGAGGGGACTGGGGGTTCCACTAGCACCCCACTTAACCTGATCA 2171  
Db 2109 ATGCAGTCTGAACAAGAGGGGACTGGGGGTTCCACTAGCACCCCACTTAACCTGATCA 2168  
Qy 2172 GCCAGGACGAGTGTGTGGGAAGGCGGCCAGCGCTGTCTCTCTCCCAATGAAGCA 2231  
Db 2169 GCCAGGACGAGTGTGTGGGAAGGCGGCCAGCGCTGTCTCTCTCCCAATGAAGCA 2228  
Qy 2232 CCATCTCACTGCTTCCAGACATTTTACCAAGTCTGCCATGAACACAGTAGGAAAGTTTA 2291  
Db 2229 CCATCTCACTGCTTCCAGACATTTTACCAAGTCTGCCATGAACACAGTAGGAAAGTTTA 2288  
Qy 2292 TCCAGTATGGCATTTCTTACAGTGGCAGAGCAGATGACAGGAAGATATCAGTCTCTAGTC 2351  
Db 2289 TCCAGTATGGCATTTCTTACAGTGGCAGAGCAGATGACAGGAAGATATCAGTCTCTAGTC 2348  
Qy 2352 TTGCTGAGCAGCAGTGGACAGAAAGCTTCCAGAACTTTGTCTTGGAGAGTGTATGAAG 2411  
Db 2349 TTGCTGAGCAGCAGTGGACAGAAAGCTTCCAGAACTTTGTCTTGGAGAGTGTATGAAG 2408  
Qy 2412 AAGATGAAGACAGTGAAGTGGGAGGAGAACAGCAGAGATTGTCTGAAGGTGAGCCAAAT 2471  
Db 2409 AAGATGAAGACAGTGAAGTGGGAGGAGAACAGCAGAGATTGTCTGAAGGTGAGCCAAAT 2468  
Qy 2472 CCAAGGAGCACAGCAGTGTATATCACTTCTTACAGAGACTCTCTGGGCTTTGCTGGAGG 2531  
Db 2469 CCAAGGAGCACAGCAGTGTATATCACTTCTTACAGAGACTCTCTGGGCTTTGCTGGAGG 2528  
Qy 2532 CCTACAGCTCTGCTGCCATCTTTGTTTCAAACTTCAGTGGTCTCTTCCAGAACCTGAGT 2591  
Db 2529 CCTACAGCTCTGCTGCCATCTTTGTTTCAAACTTCAGTGGTCTCTTCCAGAACCTGAGT 2588  
Qy 2592 ATCTCAAAAGTTGCAAAATACCTTAATAACAGAAACAGAAAGAAATGTTGCAGTATATG 2651  
Db 2589 ATCTCAAAAGTTGCAAAATACCTTAATAACAGAAACAGAAAGAAATGTTGCAGTATATG 2648  
Qy 2652 CTGAGAGTCCCATATATTTGTTGTGAAGATGCTGTGAAATGTTTAAAGGATATTTGGGG 2711  
Db 2649 CTGAGAGTCCCATATATTTGTTGTGAAGATGCTGTGAAATGTTTAAAGGATATTTGGGG 2708  
Qy 2712 TTTTCAAGGAGCACAACAGAGAGTGTCTGTTTAACTGAGCAGCAGCTTTTCTAC 2771  
Db 2709 TTTTCAAGGAGCACAACAGAGAGTGTCTGTTTAACTGAGCAGCAGCTTTTCTAC 2768  
Qy 2772 CTCAATGCAACGACAAAACCTTCTAGAAATATATTTCTGAGTCTTGTGCTGTAGGTAA 2831  
Db 2769 CTCAATGCAACGACAAAACCTTCTAGAAATATATTTCTGAGTCTTGTGCTGTAGGTAA 2828  
Qy 2832 CGTGTGGCACTGCTGGCAATGAAGGTATGAGATGAGTCTTGTAGGTACCAAGCTTCT 2891  
Db 2829 CGTGTGGCACTGCTGGCAATGAAGGTATGAGATGAGTCTTGTAGGTACCAAGCTTCT 2888  
Qy 2892 GGCTCAAGAGTTGAAGTGTCTTCCAGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCT 2951  
Db 2889 GGCTCAAGAG-TTGAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2946  
Qy 2952 CCTGGAAGACAGTCCCTCTTCCCTCCATGAGATCTGAGATCTTCCAGCTT 3003  
Db 2947 CCTGGAAGACAGTCCCTCTTCCCTCCATGAGATCTGAGATCTTCCAGCTT 2998

## RESULT 4

ABX72233

ID ABX72233 standard; cDNA; 2684 BP.

XX AC

XX ABX72233;

DT 03-JUN-2003 (first entry)  
XX Human NOVX polynucleotide #64.  
XX Human; NOVX; gene; ss; metabolic disorder; cardiomyopathy; diabetes; ASD;  
KW hypertension; congenital heart defect; aortic stenosis; valve disease;  
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;  
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;  
KW tuberosus sclerosis; scleroderma; atherosclerosis; infectious disease;  
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;  
KW Parkinson's disease; immune disorder; haematopoietic disorder;  
KW haemophilia; hypercoagulation; Crohn's disease; cancer.  
XX Homo sapiens.  
OS WO200281498-A2.  
PN 17-OCT-2002.  
XX 03-APR-2002; 2002WO-US10780.  
PR 03-APR-2001; 2001US-281086P.  
PR 03-APR-2001; 2001US-281136P.  
PR 05-APR-2001; 2001US-281863P.  
PR 06-APR-2001; 2001US-281906P.  
PR 10-APR-2001; 2001US-282020P.  
PR 10-APR-2001; 2001US-282930P.  
PR 12-APR-2001; 2001US-282934P.  
PR 12-APR-2001; 2001US-283512P.  
PR 13-APR-2001; 2001US-283710P.  
PR 17-APR-2001; 2001US-284234P.  
PR 19-APR-2001; 2001US-285325P.  
PR 20-APR-2001; 2001US-285381P.  
PR 20-APR-2001; 2001US-285609P.  
PR 23-APR-2001; 2001US-285748P.  
PR 23-APR-2001; 2001US-285890P.  
PR 24-APR-2001; 2001US-286068P.  
PR 25-APR-2001; 2001US-286292P.  
PR 27-APR-2001; 2001US-287213P.  
PR 02-MAY-2001; 2001US-288257P.  
PR 29-MAY-2001; 2001US-294164P.  
PR 30-MAY-2001; 2001US-294484P.  
PR 18-JUN-2001; 2001US-298952P.  
PR 19-JUN-2001; 2001US-299237P.  
PR 19-JUN-2001; 2001US-299237P.  
PR 12-SEP-2001; 2001US-299276P.  
PR 12-SEP-2001; 2001US-318750P.  
PR 25-SEP-2001; 2001US-324800P.  
PR 25-SEP-2001; 2001US-324802P.  
PR 27-SEP-2001; 2001US-325684P.  
PR 17-OCT-2001; 2001US-330143P.  
PR 14-NOV-2001; 2001US-332131P.  
PR 14-NOV-2001; 2001US-332240P.  
PR 14-NOV-2001; 2001US-332779P.  
PR 21-NOV-2001; 2001US-332115P.  
PR 04-DEC-2001; 2001US-337621P.  
PR 03-JAN-2002; 2002US-345783P.  
PR 16-JAN-2002; 2002US-350251P.  
PR 02-APR-2002; 2002US-0114270.

(CURA-) CURAGEN CORP.

PI Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA, Patturajan M;  
PI Liu X, Gusev VV, Li L, Vernet CM, Zerhusen BD, Gorman L;  
PI Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;  
PI Padigaru M, Shinkets RA, Gangolli EA, Taupier RJ, Casman SJ, Ji W;  
PI Anderson DM, Leite MW, Rastelli L, Edinger SR, Stone DJ;  
PI Macdougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;  
PI Ellerman K;  
XX WPI; 2003-046858/04.  
DR P-ESDB; ABU54605.  
XX New isolated NOVX polypeptide useful for treating atherosclerosis,

PT

PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,  
XX neurodegenerative disorders, Alzheimer's disease and cancer  
PS Claim 17; Page 231-232; 666pp; English.  
XX The invention relates to human polypeptides, termed NOVX, and the  
CC polynucleotides encoding them. The polypeptides and polynucleotides are  
CC useful for diagnosing disease, and screening for potential therapeutic  
CC agents. The sequences are useful for treating metabolic disorders, aortic  
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic  
CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,  
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular  
CC septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,  
CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative  
CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,  
CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease  
CC and cancer. Sequences ABX72170-ABX72275 represent human NOVX  
XX polynucleotides of the invention.  
SQ Sequence 2684 BP; 754 A; 583 C; 615 G; 732 T; 0 other;  
Query Match 86.3%; Score 2591.4; DB 25; Length 2684;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2606; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
310 AGCAGCATGTTGGGAATTACACTTTCTGACATGGATGAAATCGCACTGACCCCTTGGTAC 369  
70 AGCAGCATGTTGGGAATTACACTTTCTGACATGGATGAAATCGCACTGACCCCTTGGTAC 129  
370 AATAGATGTTTCTTATCTGCGCACATTCATCAGAAATCAGATGTTGGTGCATGTAAGCACAC 429  
130 AATAGATGTTTCTTATCTGCCACATTCATCAGAAATCAGATGTTGGTGCATGTAAGCACAC 189  
430 AAGTGGGAATGGGTGAGTGTGGCTTTAGACCCACCGTCTTCAGATCTGCAACTTTAAA 489  
190 AAGTGGGAATGGGTGAGTGTGGCTTTAGACCCACCGTCTTCAGATCTGCAACTTTAAA 249  
490 ATGGAAGAAGGCTAATGATGTCGGAAGGCAATTTGTTGGGAAGATGTTGTACTCCTG 549  
250 ATGGAAGAAGGCTAATGATGTCGGAAGGCAATTTGTTGGGAAGATGTTGTACTCCTG 309  
550 CACTCCCCAGAGCTGGGACAAAATTTTCAACCCCGATATCCCGTCTTTGGGTTTGGGAA 609  
310 CACTCCCCAGAGCTGGGACAAAATTTTCAACCCCGATATCCCGTCTTTGGGTTTGGGAA 369  
610 TGTATTATATCAATGAAGTACACACACACCGCGATGGCTTGCAGAGACGCTTTC 669  
370 TGTATTATATCAATGAAGTACACACACACCGCGATGGCTTGCAGAGACGCTTTC 429  
670 TTAGGTTCTTTTATTTCAAGAGGAGATGTCATAGGGCATGTTTGCACCAATGTGAC 729  
430 TTAGGTTCTTTTATTTCAAGAGGAGATGTCATAGGGCATGTTTGCACCAATGTGAC 489  
730 TGAATATGTCGCTCAACAGCAGTAGAGTACAGAGGCAATTTGCAGAGTGGCTGCTGAAT 789  
490 TGAATATGTCGCTCAACAGCAGTAGAGTACAGAGGCAATTTGCAGAGTGGCTGCTGAAT 549  
790 AAACCCCTGATGGTTCTGCCCGAGCAATCAAAAGCGGTTAA CAAAGTGAAGAAAGAAC 849  
550 AAACCCCTGATGGTTCTGCCCGAGCAATCAAAAGCGGTTAA CAAAGTGAAGAAAGAAC 609  
850 TAAAGGATTTCTCAAGAAATGGTGGCATGTCATACCGGCAATGATCAGATGACTGG 909  
610 TAAAGGATTTCTCAAGAAATGGTGGCATGTCATACCGGCAATGATCAGATGACTGG 669  
910 GTGGGTGCTGTAATACTGTTTCAACAGCTTTCTTTTGGAAATTTCAAAATTCACAAAGTCA 969  
670 GTGGGTGCTGTAATACTGTTTCAACAGCTTTCTTTTGGAAATTTCAAAATTCACAAAGTCA 729  
970 ACTTGAGATGGTTAAAGCTGCACTGAGCAATTTGCCGCTTCTGTTTCTACAGTTCA 1029  
730 ACTTGAGATGGTTAAAGCTGCACTGAGCAATTTGCCGCTTCTGTTTCTACAGTTCA 789

QY 1030 TAGATCCCATATTGACTATCTCTGCTCACTTCTTCAATCTTCTTCCATAACAATCAAGC 1089  
DB 790 TAGATCCCATATTGACTATCTCTGCTCACTTCTTCAATCTTCTTCCATAACAATCAAGC 849  
QY 1090 ACCATACATTGCTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTGATCCATAA 1149  
DB 850 ACCATACATTGCTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTGATCCATAA 909  
QY 1150 GCTTGGGGCTTCTTCATACGACGAAGGCTCGATGAACACACAGATGGACGGAAGATGT 1209  
DB 910 GCTTGGGGCTTCTTCATACGACGAAGGCTCGATGAACACACAGATGGACGGAAGATGT 969  
QY 1210 TCTCTATAGAGCTTGTCTCCATGGGCATATAGTTGAATTAATCTTCGACAGCAATCTT 1269  
DB 970 TCTCTATAGAGCTTGTCTCCATGGGCATATAGTTGAATTAATCTTCGACAGCAATCTT 1029  
QY 1270 GGAGATCTTCTCGAAGGACACAGTCTTAGAGTGGAAAAAACCCTTGTCTCGGGCAGG 1329  
DB 1030 GGAGATCTTCTCGAAGGACACAGTCTTAGAGTGGAAAAAACCCTTGTCTCGGGCAGG 1089  
QY 1330 ACTTTTGTGCTAGTGTAGTACTCTGCTACCAATGTCATCCAGACATCTTGATAAT 1389  
DB 1090 ACTTTTGTGCTAGTGTAGTACTCTGCTACCAATGTCATCCAGACATCTTGATAAT 1149  
QY 1390 ACCGTGTTGAAATCTCCTATGATCGCAATTCGAAGGTCACTCAATGGTGAACAACTGGG 1449  
DB 1150 ACCGTGTTGAAATCTCCTATGATCGCAATTCGAAGGTCACTCAATGGTGAACAACTGGG 1209  
QY 1450 CAAACCTAAGAGAATGAGAGCTGTGGAGTGTAGCAAGGTGTTATTAGATGTTACG 1509  
DB 1210 CAAACCTAAGAGAATGAGAGCTGTGGAGTGTAGCAAGGTGTTATTAGATGTTACG 1269  
QY 1510 AAAAACTATGTTGTGTCGAGTGGATTTTCACAGCCATTTTCTTAAAGGAATATT 1569  
DB 1270 AAAAACTATGTTGTGTCGAGTGGATTTTCACAGCCATTTTCTTAAAGGAATATT 1329  
QY 1570 AGAAAGCCAAAGTCAGAAACCGGTGCTCTCTACTTTTCCCTGGAGCAAGCGTGTGTACC 1629  
DB 1330 AGAAAGCCAAAGTCAGAAACCGGTGCTCTCTACTTTTCCCTGGAGCAAGCGTGTGTACC 1389  
QY 1630 AGCTATCTTCTTCAAGACCCAGTGTCTGTCTGTATGAAGGTGTAGAGACACGTCCTTAA 1689  
DB 1390 AGCTATCTTCTTCAAGACCCAGTGTCTGTCTGTATGAAGGTGTAGAGACACGTCCTTAA 1449  
QY 1690 TGAAGTCCGAAATGCAACAGATGAATCCCTCGAAGGAGTGTATGCAAAATCTGCGTGA 1749  
DB 1450 TGAAGTCCGAAATGCAACAGATGAATCCCTCGAAGGAGTGTATGCAAAATCTGCGTGA 1509  
QY 1750 GCATATTTCTATTACCTGTAGCAAGTCTGTGCAATTTATGTCACACACATTTGTGGCTTG 1809  
DB 1510 GCATATTTCTATTACCTGTAGCAAGTCTGTGCAATTTATGTCACACACATTTGTGGCTTG 1569  
QY 1810 CTTGCTCTCTACAGACAGGAGGAATTTGATCTCTCCCAATTTGGTGAAGACTTCTT 1869  
DB 1570 CTTGCTCTCTACAGACAGGAGGAATTTGATCTCTCCCAATTTGGTGAAGACTTCTT 1629  
QY 1870 TGTGATGAAGAGGAGTCTTGGCTGTGATTTTGACCTGGGGTCTTCAGGAATTCAGA 1929  
DB 1630 TGTGATGAAGAGGAGTCTTGGCTGTGATTTTGACCTGGGGTCTTCAGGAATTCAGA 1689  
QY 1930 AGATGTAGTAATGTCATGCCATACAGCTGTGGGAAATTTGTGCACAAATCACCCACACTAG 1989  
DB 1690 AGATGTAGTAATGTCATGCCATACAGCTGTGGGAAATTTGTGCACAAATCACCCACACTAG 1749  
QY 1990 CAGGAACGATGAGTTTTTTTATCACCCCGAGCAACATGTGCCATCAGTCTTCGAACTCAA 2049  
DB 1750 CAGGAACGATGAGTTTTTTTATCACCCCGAGCAACATGTGCCATCAGTCTTCGAACTCAA 1809  
QY 2050 CTTCTACAGCAATGGGTACTTCTATGCTTTTATCATGAGGAGGCAATCATAGCTTGCAGCCT 2109  
DB 1810 CTTCTACAGCAATGGGTACTTCTATGCTTTTATCATGAGGAGGCAATCATAGCTTGCAGCCT 1869  
QY 2110 TTATGAGTTCTGAACAAGAGGGGACTGGGGGTGTCCTAGTAGCACCCACCTAACCTGAT 2169

Db 1870 TTATCAGTCTTGAAACAAGAGGGGACTGGGGGTCCCACTAGCACCCACCTAACTCAT 1929  
Qy 2170 CAGCAGGAGCAGCTGGTGGGAAAGCGCGCAGACCTGTGCTACCTTCTCTCAATGAAGG 2229  
Db 1930 CAGCAGGAGCAGCTGGTGGGAAAGCGCGCAGACCTGTGCTACCTTCTCTCAATGAAGG 1989  
Qy 2230 CACCATCTCAGTCCCTGGCCAGACATTTTACAAAGTCTGCCATGAAACAGTAGGAAAGTT 2289  
Db 1990 CACCATCTCAGTCCCTGGCCAGACATTTTACAAAGTCTGCCATGAAACAGTAGGAAAGTT 2049  
Qy 2290 TATCCAGTATGCACTTCTACAGTGGCAGAGCAGCAGCATCACCAGNAGATATCAGTCCCTAG 2349  
Db 2050 TATCCAGTATGCACTTCTACAGTGGCAGAGCAGCAGCATCACCAGNAGATATCAGTCCCTAG 2109  
Qy 2350 TCTTCTGCTGAGCAGTGGGCAACAAGCTTCCAGAACCTTGTCTTGGAAGTGTATGA 2409  
Db 2110 TCTTCTGCTGAGCAGTGGGCAACAAGCTTCCAGAACCTTGTCTTGGAAGTGTATGA 2169  
Qy 2410 AGAAGTGAAGCAGTGTCTTGGGGAGGAAACAGCGAGATTGCTACCTGAAGGTGAGCCA 2469  
Db 2170 AGAAGTGAAGCAGTGTCTTGGGGAGGAAACAGCGAGATTGCTACCTGAAGGTGAGCCA 2229  
Qy 2470 ATCCAGGAGCAGCAGCTTATCAGCTTCTTACAGAGACTCTCTGGGCTTGTCTGGA 2529  
Db 2230 ATCCAGGAGCAGCAGCTTATCAGCTTCTTACAGAGACTCTCTGGGCTTGTCTGGA 2289  
Qy 2530 GGCCTACAGCTCTGCTGCGCATCTTGTTCACAACTTTCAGTGTCTGCTGCTCCAGAACCTGA 2589  
Db 2290 GGCCTACAGCTCTGCTGCGCATCTTGTTCACAACTTTCAGTGTCTGCTCCAGAACCTGA 2349  
Qy 2590 GTATCTGCAAAAGTTGCAAAATACCTTAATACCAAGAACAGAAAGATGTCAGATATA 2649  
Db 2350 GTATCTGCAAAAGTTGCAAAATACCTTAATACCAAGAACAGAAAGATGTCAGATATA 2409  
Qy 2650 TGCTGAGAGTGCACATATGCTTGTGAAGATGCTGTGAAATGTTTAAAGGATATGG 2709  
Db 2410 TGCTGAGAGTGCACATATGCTTGTGAAGATGCTGTGAAATGTTTAAAGGATATGG 2469  
Qy 2710 GGTCTTCAAGAGACCAACAAAGAGAGTGTCTGTTTGTAGAACCTGACAGCAGCTTTCT 2769  
Db 2470 GGTCTTCAAGAGACCAACAAAGAGAGTGTCTGTTTGTAGAACCTGACAGCAGCTTTCT 2529  
Qy 2770 ACCTCAATGCAACCCGACAAAACCTTCTAGATATATCTGAGTTTGTGCTGTAGGT 2829  
Db 2530 ACCTCAATGCAACCCGACAAAACCTTCTAGATATATCTGAGTTTGTGCTGTAGGT 2589  
Qy 2830 AACGTGTGCACTGTGCGCAATGAAGTCTATGAGATGAGTTCCTTGTAGTACAGCTT 2889  
Db 2590 AACGTGTGCACTGTGCGCAATGAAGTCTATGAGATGAGTTCCTTGTAGTACAGCTT 2649  
Qy 2890 CTGGCTCAAGAGTTTGAAGGTGCCTTCCGAGG 2922  
Db 2650 CTGGCTCAAGAG-TTGAAGGTGCCGTCGAGG 2681

RESULT 5

ID ABK94821 standard; DNA; 2487 BP.  
XX AC ABK94821;  
XX AC ABK94821;  
XX 29-AUG-2002 (first entry)  
DE DNA encoding human glycerol-3-phosphate acyltransferase hGPAT.  
XX Fatty acid regulated gene; polyunsaturated fatty acid disorder;  
KW PUFA disorder; eczema; cardiovascular disorder; hypertriglyceridaemia;  
KW dyslipidaemia; atherosclerosis; coronary artery disease;  
KW cerebrovascular disease; peripheral vascular disease; inflammation;  
KW sinusitis; asthma; pancreatitis; osteoarthritis; rheumatoid arthritis;  
KW acne; body weight disorder; obesity; cachexia; anorexia;  
KW psychiatric disorder; cancer; cystic fibrosis; pre-menstrual syndrome;

KW diabetes; diabetic complication; genetic polymorphism; gene; ds.  
XX Homo sapiens.  
OS WO200240666-A2.  
PN 23-MAY-2002.  
XX 19-NOV-2001; 2001WO-CA01632.  
XX 17-NOV-2000; 2000US-248589P.  
PR (XENO-) XENON GENETICS INC.  
XX Winther MD, Goldberg YP, Knickle LC, Haardt M, Allen SJ, Ponton A;  
PI De Antueno RJ, Jenkins DK, Nwaka SO;  
XX WPI; 2002-508327/54.  
DR P-PSDB; ABG66665.  
XX Novel isolated polypeptide segment encoded by fat regulated genes,  
PT useful for diagnosing the presence of or a predisposition for a  
PT disorder involving fatty acid regulated genes in a subject -  
XX Claim 1; Fig 28; 225pp; English.  
XX The invention describes an isolated polypeptide segment (I) whose genes  
CC are fat regulated. (I) or the polynucleotide encoding it (II) are useful  
CC for diagnosing the presence of or a predisposition for a disorder  
CC involving fatty acid regulated genes in a subject. A composition  
CC containing (I) or (II) is useful for treating a disorder involving fatty  
CC acid regulated genes, where the disorder is selected from a  
CC polyunsaturated fatty acid (PUFA) disorder, eczema, cardiovascular  
CC disorders (such as hypertriglyceridaemia, dyslipidaemia, atherosclerosis,  
CC coronary artery disease, cerebrovascular disease or peripheral vascular  
CC disease), inflammation (such as sinusitis, asthma, pancreatitis,  
CC osteoarthritis, rheumatoid arthritis or acne), body weight disorders  
CC (such as obesity, cachexia or anorexia), psychiatric disorders, cancer,  
CC cystic fibrosis, pre-menstrual syndrome, diabetes, and diabetic  
CC complications. (I) or (II) is useful as research agent and materials for  
CC discovery of treatments and diagnostics for a disease, particularly human  
CC disease. (II) is useful for constructing nucleotide probes and primers,  
CC for detecting genetic polymorphism, for detecting changes in the level of  
CC expression of (II), and as a diagnostic tool. This sequence encodes  
CC a protein regulated by fatty acids.  
XX Sequence 2487 BP; 706 A; 548 C; 570 G; 663 T; 0 other;  
SQ Query Match 82.8%; Score 2485.4; DB 24; Length 2487;  
Beat Local Similarity 100.0%; Pred. No. 0;  
Matches 2486; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 341 ATGGATGAATCTGCATCGACCTTGGTACAAATAGATTTCTTATCTGCCACATTCATCA 400  
Db 1 ATGGATGAATCTGCATCGACCTTGGTACAAATAGATTTCTTATCTGCCACATTCATCA 60  
Qy 401 GAATACAGTGTGTGTCATGTAGCACACAAAGTGGGCTGAGTGGCTTTAGA 460  
Db 61 GAATACAGTGTGTGTCATGTAGCACACAAAGTGGGCTGAGTGGCTTTAGA 120  
Qy 461 CCCACCTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTTAATGAGTCGGAAGG 520  
Db 121 CCCACCATCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTTAATGAGTCGGAAGG 180  
Qy 521 CCATTTGTTGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGCAAAATTTTCAAC 580  
Db 181 CCATTTGTTGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGCAAAATTTTCAAC 240  
Qy 581 CCCAGTATCCGCTTTGGGTTGGCGAATGTTATTTATATCAATGAATCAACACAGA 640  
Db 241 CCCAGTATCCGCTTTGGGTTGGCGAATGTTATTTATATCAATGAATCAACACAGA 300  
Qy 641 CACGCGGATGGCTTCAAGAGCGCTTCTTACGTTCTTTTATTCAAGAGCGAGATGTG 700

Db 301 |||||CACCGGATGGCTTGCAGACGCTTTCTTACGTTCTTTTATCAAGACGAGATGTG 360  
Qy 701 |||||CATAGGGCATGTTTCCCAACATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 760  
Db 361 |||||CATAGGGCATGTTTCCCAACATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 420  
Qy 761 |||||GAGCAATTCAGAAAGTGGCTGTGTAATTAACCCCTGATGGTTCTGCCAGCAGCAATCA 820  
Db 421 |||||GAGCAATTCAGAAAGTGGCTGTGTAATTAACCCCTGATGGTTCTGCCAGCAGCAATCA 480  
Qy 821 |||||AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTAAAGGATCTTCAAGAAATGGTTGCCACT 880  
Db 481 |||||AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTAAAGGATCTTCAAGAAATGGTTGCCACT 540  
Qy 881 |||||GTCTCACCGGCAATGATCAGACTGACTGGTGGGTGCTGTAATTAACCTGTTCAACAGCTTC 940  
Db 541 |||||GTCTCACCGGCAATGATCAGACTGACTGGTGGGTGCTGTAATTAACCTGTTCAACAGCTTC 600  
Qy 941 |||||TTTTGGAACATTCAAAATTCACAAAGGTCAACTTTGAGATGGTTAAAGCTGCAACTGAGACG 1000  
Db 601 |||||TTTTGGAACATTCAAAATTCACAAAGGTCAACTTTGAGATGGTTAAAGCTGCAACTGAGACG 660  
Qy 1001 |||||AATTGCGGCTTCTGTTTCTACAGTTTCATAGATCCCATATTCATATTCGCTGCTCACT 1060  
Db 661 |||||AATTGCGGCTTCTGTTTCTACAGTTTCATAGATCCCATATTCATATTCGCTGCTCACT 720  
Qy 1061 |||||TTCAATCTCTCTGCGCATACATCAAGACCAATACATTTGCTTCAAGCAATATCTCAAC 1120  
Db 721 |||||TTCAATCTCTCTGCGCATACATCAAGACCAATACATTTGCTTCAAGCAATATCTCAAC 780  
Qy 1121 |||||ATCCCAATCTTCAGTACCTTTGATCCATAAGCTTTGGGGCTTCTTATACAGCAAGGCTC 1180  
Db 781 |||||ATCCCAATCTTCAGTACCTTTGATCCATAAGCTTTGGGGCTTCTTATACAGCAAGGCTC 840  
Qy 1181 |||||GATGAACACACAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1240  
Db 841 |||||GATGAACACACAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Qy 1241 |||||GTTGAATTAATTCGACAGCAGCAATCTTGGAGATCTTCTGGAAGGCAACGCTTCTAGG 1300  
Db 901 |||||GTTGAATTAATTCGACAGCAGCAATCTTGGAGATCTTCTGGAAGGCAACGCTTCTAGG 960  
Qy 1301 |||||AGTGGAAAAACCTCTTGTGCTCGGGCAGGACTTTTGTGAGTGTGGTAGATCTCTGTCT 1360  
Db 961 |||||AGTGGAAAAACCTCTTGTGCTCGGGCAGGACTTTTGTGAGTGTGGTAGATCTCTGTCT 1020  
Qy 1361 |||||ACCAATGTCTATCCAGACATCTTTGATAATACCTGTTGGAAATCTCTATGATCGCATATC 1420  
Db 1021 |||||ACCAATGTCTATCCAGACATCTTTGATAATACCTGTTGGAAATCTCTATGATCGCATATC 1080  
Qy 1421 |||||GAAGGTCACTACATGGTGAACAACTGGGCAACCTTAAGAAATGAGAGCTGTGGAGT 1480  
Db 1081 |||||GAAGGTCACTACATGGTGAACAACTGGGCAACCTTAAGAAATGAGAGCTGTGGAGT 1140  
Qy 1481 |||||GTAGCAAGAGGTGTATTAAGAAATGTTACGAAAAAATATGTTGTGTCCGAGTGGATTTT 1540  
Db 1141 |||||GTAGCAAGAGGTGTATTAAGAAATGTTACGAAAAAATATGTTGTGTCCGAGTGGATTTT 1200  
Qy 1541 |||||GCAAGCCATTTTCTTTAAAGGAATATTTAAGAAAGCCAAAGTCCAGAAAACCGGTGTCTGCT 1600  
Db 1201 |||||GCAAGCCATTTTCTTTAAAGGAATATTTAAGAAAGCCAAAGTCCAGAAAACCGGTGTCTGCT 1260  
Qy 1601 |||||CTACTTTCCTGAGCAGAGGTTGTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1660  
Db 1261 |||||CTACTTTCCTGAGCAGAGGTTGTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320  
Qy 1661 |||||GCTGATGAAGGTAGAGACAGTCCATTAAAGTCCAGAAATGCAACAGATGAATCCCTA 1720  
Db 1321 |||||GCTGATGAAGGTAGAGACAGTCCATTAAAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Qy 1721 |||||CGAAGGAGGTGTGATTCGAAATCTGGCTGAGCATATTCATTCATCTGCTAGCAAGTCTGCT 1780

Db 1381 |||||CGAAGGAGGTGTGATTCGAAATCTGGCTGAGCATATTTCTATTCTACTCTAGCAAGTCTGCT 1440  
Qy 1781 |||||GCCATTATGTCCACACACATTTGGCTTGCCTCTCTCTACAGACACAGGACGGAAT 1840  
Db 1441 |||||GCCATTATGTCCACACACATTTGGCTTGCCTCTCTCTACAGACACAGGACGGAAT 1500  
Qy 1841 |||||GATCTCTCCACATTTGGTCCGAAGACTTTCTTTGTGATGAAGAGGAAGTCTCTGGCTCGTGAT 1900  
Db 1501 |||||GATCTCTCCACATTTGGTCCGAAGACTTTCTTTGTGATGAAGAGGAAGTCTCTGGCTCGTGAT 1560  
Qy 1901 |||||TTTGACCTGGGTTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGCTGCTG 1960  
Db 1561 |||||TTTGACCTGGGTTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGCTGCTG 1620  
Qy 1961 |||||GGAAATTTGTGCACAAATCACCCACACTAGCAGGAAGTGTAGTTTTTTTATCACCCCGAC 2020  
Db 1621 |||||GGAAATTTGTGCACAAATCACCCACACTAGCAGGAAGTGTAGTTTTTTTATCACCCCGAC 1680  
Qy 2021 |||||ACAACTGTCCCATCAGTCTTTCGAACTCAACTTCTACAGCAATGGGTACTTCAATGCTTT 2080  
Db 1681 |||||ACAACTGTCCCATCAGTCTTTCGAACTCAACTTCTACAGCAATGGGTACTTCAATGCTTT 1740  
Qy 2081 |||||ATCATGGAGGCCATCATAGCTTTCAGCTTTATGAGTTCTGNAAGAAGGCACTGGGG 2140  
Db 1741 |||||ATCATGGAGGCCATCATAGCTTTCAGCTTTATGAGCTTTCTGAAACAAGAGGGACTGGGG 1800  
Qy 2141 |||||GGTCCCACTAGCACCCCACTAACCTGTAGCAGCAGGAGCAGCTGGTGGGAAGCGGCC 2200  
Db 1801 |||||GGTCCCACTAGCACCCCACTAACCTGTAGCAGCAGGAGCAGCTGGTGGGAAGCGGCC 1860  
Qy 2201 |||||AGCCTGTGCTACCTTCTCTCAATGAAGCAACCATCTCACTGCTTTCGAGACATTTTAC 2260  
Db 1861 |||||AGCCTGTGCTACCTTCTCTCAATGAAGCAACCATCTCACTGCTTTCGAGACATTTTAC 1920  
Qy 2261 |||||CAAGTCTGCCATGAACACAGTAGAAAGTTTATCCAGTATGGCAATCTTACAGTGGCAGAG 2320  
Db 1921 |||||CAAGTCTGCCATGAACACAGTAGAAAGTTTATCCAGTATGGCAATCTTACAGTGGCAGAG 1980  
Qy 2321 |||||CAGATGACCCAGGAGATATCAGTCTTGTGAGCAGCAGTGGGACAGAAAGCTT 2380  
Db 1981 |||||CAGATGACCCAGGAGATATCAGTCTTGTGAGGAGCTTACAGTCTGCTGCCATCTTTTTCAC 2040  
Qy 2381 |||||CCAGAACCTTTGTCTTGGAGAAAGTGTATGAAGAGATGAAGACAGTCACTTTGGGGAGGAA 2440  
Db 2041 |||||CCAGAACCTTTGTCTTGGAGAAAGTGTATGAAGAGATGAAGACAGTCACTTTGGGGAGGAA 2100  
Qy 2441 |||||CAGCGAGATTTGCTACTGAAGGTGAGCCAAATCCAGGAGGACAGCAGTTCATCCTTC 2500  
Db 2101 |||||CAGCGAGATTTGCTACTGAAGGTGAGCCAAATCCAGGAGGACAGCAGTTCATCCTTC 2160  
Qy 2501 |||||TTACAGAGACTCTTGGGCTTTTGTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2560  
Db 2161 |||||TTACAGAGACTCTTGGGCTTTTGTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2220  
Qy 2561 |||||AACTTCAGTGGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACTTAATA 2620  
Db 2221 |||||AACTTCAGTGGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACTTAATA 2280  
Qy 2621 |||||ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTCCACATATTTGTCTTGTGAAG 2680  
Db 2281 |||||ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTCCACATATTTGTCTTGTGAAG 2340  
Qy 2681 |||||AATGCTGTCAAAATGTTTAAAGGATATTTGGGTTTTTCAAGAGACCAAAACAAAGAGAGTG 2740  
Db 2341 |||||AATGCTGTCAAAATGTTTAAAGGATATTTGGGTTTTTCAAGAGACCAAAACAAAGAGAGTG 2400  
Qy 2741 |||||TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCCGACAAACCTTCTAGAA 2800  
Db 2401 |||||TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCCGACAAACCTTCTAGAA 2460  
Qy 2801 |||||TATATTTCTGAGTTTGTGCTGTAG 2827  
Db 2461 |||||TATATTTCTGAGTTTGTGCTGTAG 2487



RESULT 6  
AAL55475  
ID AAL55475 standard; DNA; 2487 BP.  
XX  
AC AAL55475;  
XX  
DT 22-MAY-2003 (first entry)  
XX  
DE GPAM related DNA sequence, SEQ ID No 1.  
XX  
KW Antidiabetic; nephrotropic; neuroprotective; ophthalmological; human;  
KW mitochondrial sn-glycerol-3-phosphate acyltransferase; GPAM;  
KW diabetic complication; retinopathy; neuropathy; enzyme; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 1..2487  
FT /\*tag= a  
ET /product= "GPAM-related protein"  
XX  
PN WO2003008590-A1.  
XX  
PD 30-JAN-2003.  
XX  
PF 16-JUL-2002; 2002WO-JP07189.  
XX  
PR 16-JUL-2001; 2001JP-0215337.  
XX  
PA (KISP) KISSEI PHARM CO LTD.  
XX  
PI Sakamoto S, Onota H, Sugano S, Nakamura Y;  
XX  
DR WPI; 2003-229583/22.  
DR P-PSDB; AAO27073.  
XX  
PT Human mitochondrial sn-glycerol-3-phosphate acyltransferase and  
PT antagonists for treatment and prevention of diabetic complications  
XX  
PS Claim 4; Page 33-40; 56pp; Japanese.  
XX  
CC The invention relates to a novel protein having human mitochondrial sn-  
CC glycerol-3-phosphate acyltransferase (GPAM) activity. The novel protein  
CC with GPAM activity can be used in the prevention and treatment of  
CC diabetic complications, including retinopathy and neuropathy, by  
CC administration of antagonists to human GPAM. This polynucleotide sequence  
CC represents a DNA sequence encoding a protein relating to the GPAM  
CC activity protein of the invention.  
XX  
SQ Sequence 2487 BP; 704 A; 549 C; 571 G; 663 T; 0 other;

Query Match 82.7%; Score 2483.8; DB 25; Length 2487;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2485; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 341 ATGGATGAATCTGCACCTGACCTGGTACATAGATGTTCTTATCTGCCACATTCATCA 400  
Db 1 ATGGATGAATCTGCACCTGACCTGGTACATAGATGTTCTTATCTGCCACATTCATCA 60  
Qy 401 GAATACAGTGTGGTCGATGTAAGCACACAAGTGAAGGAATGGGTGAGTGGCTTTTGA 460  
Db 61 GAATACAGTGTGGTCGATGTAAGCACACAAGTGAAGGAATGGGTGAGTGGCTTTTGA 120  
Qy 461 CCCACCGTCTTCAGATCTGCACTTTAAATGGAAAGAAAGCCTAATGAGTCGAAAGG 520  
Db 121 CCCACCGTCTTCAGATCTGCACTTTAAATGGAAAGAAAGCCTAATGAGTCGAAAGG 180  
Qy 521 CCATTTGTGAAGATGTTGTTACTCTCGACTCCCGAGAGCTGGGACAAATTTTCAAC 580  
Db 181 CCATTTGTGAAGATGTTGTTACTCTCGACTCCCGAGAGCTGGGACAAATTTTCAAC 240

Qy 581 CCCAGTATCCCGTCTTTGGGTTTCGGAAATGTTATATATATCAATGAACACTCACACAAGA 640  
Db 241 CCCAGTATCCCGTCTTTGGGTTTCGGAAATGTTATATATATCAATGAACACTCACACAAGA 300  
Qy 641 CACCGGATGGCTTGAAGACGCTTTCTTACGTTCTTTTATTATTCAAGACGAGATGTG 700  
Db 301 CACCGGATGGCTTGAAGACGCTTTCTTACGTTCTTTTATTATTCAAGACGAGATGTG 360  
Qy 701 CATAAAGGCGATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 760  
Db 361 CATAAAGGCGATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 420  
Qy 761 GAGCAATTCAGAAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCACGACGAAATCA 820  
Db 421 GAGCAATTCAGAAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCACGACGAAATCA 480  
Qy 821 AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTTAAAGGATTTCTTCAAGAAATGGTTGCCACT 880  
Db 481 AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTTAAAGGATTTCTTCAAGAAATGGTTGCCACT 540  
Qy 881 GTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGCTTAAACCTGTTTCAACAGCTTC 940  
Db 541 GTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGCTTAAACCTGTTTCAACAGCTTC 600  
Qy 941 TTTTGAACATTCAAATTCACAAAGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACG 1000  
Db 601 TTTTGAACATTCAAATTCACAAAGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACG 660  
Qy 1001 AATTGGCGTTCGTTTCTTACAGTTTATAGATCCCATATTTGACTATTTGCTGCTCACT 1060  
Db 661 AATTGGCGTTCGTTTCTTACAGTTTATAGATCCCATATTTGACTATTTGCTGCTCACT 720  
Qy 1061 TTCAATCTCTTCTGCAATACATCAAGACCAATACATTTGCTTCCAGCAATTAATCTCAAC 1120  
Db 721 TTCAATCTCTTCTGCAATACATCAAGACCAATACATTTGCTTCCAGCAATTAATCTCAAC 780  
Qy 1121 ATCCCAATCTTCACTACCTTGATCCATAAGCTTGGGGCTTCTTTCATACGACGAGGCTC 1180  
Db 781 ATCCCAATCTTCACTACCTTGATCCATAAGCTTGGGGCTTCTTTCATACGACGAGGCTC 840  
Qy 1181 GATGAACACACAGATGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1240  
Db 841 GATGAACACACAGATGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Qy 1241 GTTGAATTTACTTGCACAGCAGCAATTTCTGGAGATCTTCTGGAGGACGACGCTTCAGG 1300  
Db 901 GTTGAATTTACTTGCACAGCAGCAATTTCTGGAGATCTTCTGGAGGACGACGCTTCAGG 960  
Qy 1301 AGTGGAAAAACCTCTTGTGCTCGGGCAGGACTTTGTGTCAGTTGTGTAGATACTCTGTCT 1360  
Db 961 AGTGGAAAAACCTCTTGTGCTCGGGCAGGACTTTGTGTCAGTTGTGTAGATACTCTGTCT 1020  
Qy 1361 ACCAATGTCATCCAGACATCTTTGATAATACTGTGTGGAATCTCCTATGATCGCAATTATC 1420  
Db 1021 ACCAATGTCATCCAGACATCTTTGATAATACTGTGTGGAATCTCCTATGATCGCAATTATC 1080  
Qy 1421 GAAGTCACTACATGCTGAACTGGGCAACCTTAAGAAAGATGAGAGCTGTGGAGT 1480  
Db 1081 GAAGTCACTACATGCTGAACTGGGCAACCTTAAGAAAGATGAGAGCTGTGGAGT 1140  
Qy 1481 GTAGCAAGAGTGTATTAGTAATGTTACGAAAAAACTATGTTGTGTCGAGTGGATTTT 1540  
Db 1141 GTAGCAAGAGTGTATTAGTAATGTTACGAAAAAACTATGTTGTGTCGAGTGGATTTT 1200  
Qy 1541 GCACAGCCATTTCTTTAAAGGAATATTAGAAAGCCAAAGTTCAGAAACCGGTGTCTGCT 1600  
Db 1201 GCACAGCCATTTCTTTAAAGGAATATTAGAAAGCCAAAGTTCAGAAACCGGTGTCTGCT 1260  
Qy 1601 CTACTTTTCCCTGGAGCAGGTTGTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1660  
Db 1261 CTACTTTTCCCTGGAGCAGGTTGTTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320  
Qy 1661 GCTGATGAAGTAGAGACACGTCCTCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1720



Best Local Similarity 99.6%; Pred. No. 0; Matches 2359; Conservative 0; Mismatches 7; Indels 2; Gaps 2;											
QY	636	CAAGACACCGCGATGGCTTGCAAGACGCGCTTCTTACGTTCTTTTATTTCAAGAGCGAG	695	QY	1716	CCCTACGAAGGAGGTTCATTGCAAAATCTGGCTGAGCATATTTCTATTCACCTGCTAGCAAGT	1775	QY	1716	CCCTACGAAGGAGGTTCATTGCAAAATCTGGCTGAGCATATTTCTATTCACCTGCTAGCAAGT	1775
Db	173	CCAGACACCGCGATGGCTTGCAAGACGCGCTTCTTACGTTCTTTTATTTCAAGAGCGAG	232	Db	1253	CCCTACGAAGGAGGTTCATTGCAAAATCTGGCTGAGCATATTTCTATTCACCTGCTAGCAAGT	1312	Db	1253	CCCTACGAAGGAGGTTCATTGCAAAATCTGGCTGAGCATATTTCTATTCACCTGCTAGCAAGT	1312
QY	696	ATGTGCATAAGGCGATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAG	755	QY	1776	CCTGTGCCATTTATGTCACACACATTTGTGGCTTGCCCTGCTCTCTACAGACACAGGCGAG	1835	QY	1776	CCTGTGCCATTTATGTCACACACATTTGTGGCTTGCCCTGCTCTCTACAGACACAGGCGAG	1835
Db	233	ATGTGCATAAGGCGATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAG	292	Db	1313	CCTGTGCCATTTATGTCACACACATTTGTGGCTTGCCCTGCTCTCTACAGACACAGGCGAG	1372	Db	1313	CCTGTGCCATTTATGTCACACACATTTGTGGCTTGCCCTGCTCTCTACAGACACAGGCGAG	1372
QY	756	TACAAGAGGCAATTGCAGAAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCAGCAGC	815	QY	1836	GAATTTGATCTCTCCACATTTGTCGAAGACTTCTTTTGTGATGAAAGAGGAAGTCTCTGGCTC	1895	QY	1836	GAATTTGATCTCTCCACATTTGTCGAAGACTTCTTTTGTGATGAAAGAGGAAGTCTCTGGCTC	1895
Db	293	TACAAGAGGCAATTGCAGAAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCAGCAGC	352	Db	1373	GAATTTGATCTCTCCACATTTGTCGAAGACTTCTTTTGTGATGAAAGAGGAAGTCTCTGGCTC	1432	Db	1373	GAATTTGATCTCTCCACATTTGTCGAAGACTTCTTTTGTGATGAAAGAGGAAGTCTCTGGCTC	1432
QY	816	AATCAAAAGCCGTTAAACAAAGTGAAGAAAGCTAAAGAGGATCTTCAAGAAATGGTTG	875	QY	1896	GTGATTTTTCACCTCGGGTTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGC	1955	QY	1896	GTGATTTTTCACCTCGGGTTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGC	1955
Db	353	AATCAAAAGCCGTTAAACAAAGTGAAGAAAGCTAAAGAGGATCTTCAAGAAATGGTTG	412	Db	1433	GTGATTTTTCACCTCGGGTTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGC	1492	Db	1433	GTGATTTTTCACCTCGGGTTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGC	1492
QY	876	CCACTGTCTCACCGCAATGATCAGACTGACTGGGTGGTGTCTGCTAAAACCTGTTCAACA	935	QY	1956	TGCTGGGAAATTTGTGTCAATCACCCACACTAGCAGGAACAGTAGAGTTTTTTTATCACCC	2015	QY	1956	TGCTGGGAAATTTGTGTCAATCACCCACACTAGCAGGAACAGTAGAGTTTTTTTATCACCC	2015
Db	413	CCACTGTCTCACCGCAATGATCAGACTGACTGGGTGGTGTCTGCTAAAACCTGTTCAACA	472	Db	1493	TGCTGGGAAATTTGTGTCAATCACCCACACTAGCAGGAACAGTAGAGTTTTTTTATCACCC	1552	Db	1493	TGCTGGGAAATTTGTGTCAATCACCCACACTAGCAGGAACAGTAGAGTTTTTTTATCACCC	1552
QY	936	GCTTCTTTTGGAACTTCAAATTCACAAAGGTCAAATTCAGATGTTTAAAGCTGCAACTG	995	QY	2016	CCAGCACAACTGTCCTCAGTCAGTCTTCGAACTCAACTTCTACAGCAATGGGGTACTTCATG	2075	QY	2016	CCAGCACAACTGTCCTCAGTCAGTCTTCGAACTCAACTTCTACAGCAATGGGGTACTTCATG	2075
Db	473	GCTTCTTTTGGAACTTCAAATTCACAAAGGTCAAATTCAGATGTTTAAAGCTGCAACTG	532	Db	1553	CCAGCACAACTGTCCTCAGTCAGTCTTCGAACTCAACTTCTACAGCAATGGGGTACTTCATG	1612	Db	1553	CCAGCACAACTGTCCTCAGTCAGTCTTCGAACTCAACTTCTACAGCAATGGGGTACTTCATG	1612
QY	996	AGACGAAATTTGGCGCTTCTGTTTCTACCAAGTTTCATAGATCCCATATTCAGTCTGCTGC	1055	QY	2076	TCTTTTATCATGAGGCCATCATAGCTTGCAGCCTTTTATGCAGTTCTCAAAAGAGGGGAC	2135	QY	2076	TCTTTTATCATGAGGCCATCATAGCTTGCAGCCTTTTATGCAGTTCTCAAAAGAGGGGAC	2135
Db	533	AGACGAAATTTGGCGCTTCTGTTTCTACCAAGTTTCATAGATCCCATATTCAGTCTGCTGC	592	Db	1613	TCTTTTATCATGAGGCCATCATAGCTTGCAGCCTTTTATGCAGTTCTCAAAAGAGGGGAC	1672	Db	1613	TCTTTTATCATGAGGCCATCATAGCTTGCAGCCTTTTATGCAGTTCTCAAAAGAGGGGAC	1672
QY	1056	TCACATTTCTCTTCTGCCATAACATCAAGACCACTACATTCCTTTCAGGCAATAATC	1115	QY	2136	TGGGGGTGCCACTAGCACCCCACTTAACCTGTATCAGCCAGGAGCAGCTGGTGGGAAGG	2195	QY	2136	TGGGGGTGCCACTAGCACCCCACTTAACCTGTATCAGCCAGGAGCAGCTGGTGGGAAGG	2195
Db	593	TCACATTTCTCTTCTGCCATAACATCAAGACCACTACATTCCTTTCAGGCAATAATC	652	Db	1673	TGGGGGTGCCACTAGCACCCCACTTAACCTGTATCAGCCAGGAGCAGCTGGTGGGAAGG	1732	Db	1673	TGGGGGTGCCACTAGCACCCCACTTAACCTGTATCAGCCAGGAGCAGCTGGTGGGAAGG	1732
QY	1116	TCACATTTCTCTTCTGCCATAACATCAAGACCACTACATTCCTTTCAGGCAATAATC	1175	QY	2196	CGGCCAGAGCTGTGTACCTCTCTCCAAATGAAGCAGCATCTCACTGCCTTGCAGACAT	2255	QY	2196	CGGCCAGAGCTGTGTACCTCTCTCCAAATGAAGCAGCATCTCACTGCCTTGCAGACAT	2255
Db	653	TCACATTTCTCTTCTGCCATAACATCAAGACCACTACATTCCTTTCAGGCAATAATC	712	Db	1733	CGGCCAGAGCTGTGTACCTCTCTCCAAATGAAGCAGCATCTCACTGCCTTGCAGACAT	1792	Db	1733	CGGCCAGAGCTGTGTACCTCTCTCCAAATGAAGCAGCATCTCACTGCCTTGCAGACAT	1792
QY	1176	GGCTCGATGAACACACAGATGGAAGGAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGC	1235	QY	2256	TTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTAGGCAATCTTACAGTGG	2315	QY	2256	TTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTAGGCAATCTTACAGTGG	2315
Db	713	GGCTCGATGAACACACAGATGGAAGGAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGC	772	Db	1793	TTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTAGGCAATCTTACAGTGG	1852	Db	1793	TTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTAGGCAATCTTACAGTGG	1852
QY	1236	ATATAGTTGAATTTACTTTCAGACAGCAGCAATTTCTGGAGATCTTCTGGAAGGCAACAGTT	1295	QY	2316	CAGACAGCATGACACAGGAAGATATCAGTCTTAGTCTTGTGTGAGCAGCAGTGGGCAAGA	2375	QY	2316	CAGACAGCATGACACAGGAAGATATCAGTCTTAGTCTTGTGTGAGCAGCAGTGGGCAAGA	2375
Db	773	ATATAGTTGAATTTACTTTCAGACAGCAGCAATTTCTGGAGATCTTCTGGAAGGCAACAGTT	832	Db	1853	CAGACAGCATGACACAGGAAGATATCAGTCTTAGTCTTGTGTGAGCAGCAGTGGGCAAGA	1912	Db	1853	CAGACAGCATGACACAGGAAGATATCAGTCTTAGTCTTGTGTGAGCAGCAGTGGGCAAGA	1912
QY	1296	CTAGGAGTGGAAAACTTCTGCTCGGCGAGGACTTTTGTTCAGTTGTGTGTAGATCTC	1355	QY	2376	AGCTTTCAGAACCTTTGCTTGGAGAGTGTATGAAGAGATGAAGACAGTACTTTGGGG	2435	QY	2376	AGCTTTCAGAACCTTTGCTTGGAGAGTGTATGAAGAGATGAAGACAGTACTTTGGGG	2435
Db	833	CTAGGAGTGGAAAACTTCTGCTCGGCGAGGACTTTTGTTCAGTTGTGTGTAGATCTC	892	Db	1913	AGCTTTCCTGAAACCTTTGCTTGGAGAGTGTATGAAGAGATGAAGACAGTACTTTGGGG	1972	Db	1913	AGCTTTCCTGAAACCTTTGCTTGGAGAGTGTATGAAGAGATGAAGACAGTACTTTGGGG	1972
QY	1356	TGCTTACCAATGTCTCCAGACATCTTGATAATACCTGTTGGATCTCTCTATGATCGCA	1415	QY	2436	AGGAAACAGCAGATTTGCTTACCTGAAGGTGAGCCAAATCCAAAGGAGCACACAGTTTATCA	2495	QY	2436	AGGAAACAGCAGATTTGCTTACCTGAAGGTGAGCCAAATCCAAAGGAGCACACAGTTTATCA	2495
Db	893	TGCTTACCAATGTCTCCAGACATCTTGATAATACCTGTTGGAAATCTCTTATGATCGCA	952	Db	1973	AGGAAACAGCAGATTTGCTTACCTGAAGGTGAGCCAAATCCAAAGGAGCACACAGTTTATCA	2032	Db	1973	AGGAAACAGCAGATTTGCTTACCTGAAGGTGAGCCAAATCCAAAGGAGCACACAGTTTATCA	2032
QY	1416	TTATCGAAGGTCACTACAAATGTGTAACCACTGGGCAAACTTAAGAAAGATGAGAGCCTGT	1475	QY	2496	CCTTCTTACAGAGACTCCTTGGGCTTTGCTGGAGGCTTACAGCTCTGTGCTGCCATCTTTG	2555	QY	2496	CCTTCTTACAGAGACTCCTTGGGCTTTGCTGGAGGCTTACAGCTCTGTGCTGCCATCTTTG	2555
Db	953	TTATCGAAGGTCACTACAAATGTGTAACCACTGGGCAAACTTAAGAAAGATGAGAGCCTGT	1012	Db	2033	CCTTCTTACAGAGACTCCTTGGGCTTTGCTGGAGGCTTACAGCTCTGTGCTGCCATCTTTG	2092	Db	2033	CCTTCTTACAGAGACTCCTTGGGCTTTGCTGGAGGCTTACAGCTCTGTGCTGCCATCTTTG	2092
QY	1476	GGAGTGTAGCAAGAGGTGTTATTAGAAATGTTACGAAAAAATCATGGTTGTGTCGAGTGG	1535	QY	2556	TTCAACAACTTCAGTGGTCTGTTCCAGAACCTTGATATCTGCAAAAAGTTGCACAAAATACC	2615	QY	2556	TTCAACAACTTCAGTGGTCTGTTCCAGAACCTTGATATCTGCAAAAAGTTGCACAAAATACC	2615
Db	1013	GGAGTGTAGCAAGAGGTGTTATTAGAAATGTTTACGAAAAAATCATGGTTGTGTCGAGTGG	1072	Db	2093	TTCAACAACTTCAGTGGTCTGTTCCAGAACCTTGATATCTGCAAAAAGTTGCACAAAATACC	2152	Db	2093	TTCAACAACTTCAGTGGTCTGTTCCAGAACCTTGATATCTGCAAAAAGTTGCACAAAATACC	2152
QY	1536	ATTTTGCACAGCCATTTTCTTAAAGGAATATTTTAAAGCCAAAGTCAGAAACCGGTGT	1595	QY	2616	TAATAACACAGAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCACATATTTGCTTG	2675	QY	2616	TAATAACACAGAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCACATATTTGCTTG	2675
Db	1073	ATTTTGCACAGCCATTTTCTTAAAGGAATATTTTAAAGCCAAAGTCAGAAACCGGTGT	1132	Db	2153	TAATAACACAGAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCACATATTTGCTTG	2212	Db	2153	TAATAACACAGAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCACATATTTGCTTG	2212
QY	1596	CTGCTCTACTTTCCCTGGAGCAAGCTTGTGTACCAAGCTATATCTTCCCTCAAGACCCAGTG	1655	QY	2676	TGAAGAAATGCTGTGAAAAATTTTAAAGGATATTTGGGGTTTTTCAAGGAGACCAAAACAAA	2735	QY	2676	TGAAGAAATGCTGTGAAAAATTTTAAAGGATATTTGGGGTTTTTCAAGGAGACCAAAACAAA	2735
Db	1133	CTGCTCTACTTTCCCTGGAGCAAGCTTGTGTACCAAGCTATATCTTCCCTCAAGACCCAGTG	1192	Db	2213	TGAAGAAATGCTGTGAAAAATTTTAAAGGATATTTGGGGTTTTTCAAGGAGACCAAAACAAA	2272	Db	2213	TGAAGAAATGCTGTGAAAAATTTTAAAGGATATTTGGGGTTTTTCAAGGAGACCAAAACAAA	2272
QY	1656	ATGCTGCTGATGAAGGTGAGACACGCTCAATTAATGAGTCCAGAAATGCAACAGATGAAT	1715	QY	2736	GAGTGTCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAATGCAACCGACAAACCTTC	2795	QY	2736	GAGTGTCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAATGCAACCGACAAACCTTC	2795
Db	1715	ATGCTGCTGATGAAGGTGAGACACGCTCAATTAATGAGTCCAGAAATGCAACAGATGAAT	1715	Db	2273	GAGTGTCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAATGCAACCGACAAACCTTC	2332	Db	2273	GAGTGTCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAATGCAACCGACAAACCTTC	2332



Qy	1039	TATTGACGTATCTGCTGCTCACTTTTCATTCTCTTCTGCCATAACATCAAAGCACCATACAT	1098
Db	858	CATCGACTACTCTGCTGCTCACTTTTCATCTCTCTGCGCAACAATCAAAGCTCCATACAT	917
Qy	1099	TGCTTTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTTGATCCATAAGCTTTGGGG	1158
Db	918	CGCTTCGGGCACAACTCTAACATCCCAATCTTCAGTACCTTGATTCAAAGCTTTGGGG	977
Qy	1159	CTTCTTCATACGACGAAGGCTCGATGAAACACCAAGATGGACGGAAGAATGTTCTCTATAG	1218
Db	978	CTTTTTTCATAGACGGAAGGCTTGACGAAACTCCAGATGGACGCAAGACATTTCTGTACAG	1037
Qy	1219	AGCTTTGCTCCATCGGCATATAGTTGAATTAATCTTCGACAGACGAATTCCTTGAGATCTT	1278
Db	1038	AGCGTTCCTCATGGGCATATAGTTGAATTCCTTCGACAGACGAGTTCCTTGAGATCTT	1097
Qy	1279	CCTGGAAGGCACACGTTCTTAGGAGTGGAAAAACCTCTTGTCCTCGGCGACGACTTTTGTC	1338
Db	1098	CCTGGAAGGCACCCGCTCCCGCAGTGGCAAGACCTCTCTGTGCCGGCGCTCTGTC	1157
Qy	1339	AGTTGTGGTATGACTCTGTCTTCAACATGTCATCCAGACATCTTGATAATACTGTGTGG	1398
Db	1158	AGTGSTAGTGGATCTCTGTTCATCCAAACCATCCCTGTCATCTGTGTCATCCCTGTGGG	1217
Qy	1399	AATCTCTATGATCGCATATCGAAGTCACTACAATGGTGAACTGGCGCAACCTTAA	1458
Db	1218	CATCTGTATGATCGGATAATCGAAGTCACTACAATGGTGAACTGGCGCAAGGCCAA	1277
Qy	1459	GAAGAAATGAGAGCCTGTGAGTGTAGCAAGAGGTGTTATTAGAAATGTTCGAAAAACTA	1518
Db	1278	GAAGAAATGAAGTCTCTGGAGTGGCGAAGAGGCTTATCAGAATGCTCGGAANAACCTA	1337
Qy	1519	TGTTGTGTCGAGTGGATTTTGACAGCCATTTTCTTTAAAGAAATTTTGAAGGCCA	1578
Db	1338	CGGCTATGTCCGAGTGGACTTTGACAGCCATTTTCTTTCAAGGAATTTTGAAGGCCA	1397
Qy	1579	AAGTCAGAAAACCGGTGTCGTCTACTTTCCCTGGAGCAAGCGTTGTTACCACTATCT	1638
Db	1398	AAGTCAGAAAACCTGTATCTGCTCCCTCTCTTTTGGACAGACACTGTTTACGAGATCT	1457
Qy	1639	TCCTTTCAAGACCCAGTGATGCTGTGTGAAGGTAGACACGCTCCATTAATCAGTCCAG	1698
Db	1458	TCCTTTCAAGACCTGTATGCTGCTGCGGAACATGAAGACATGTCAGTAATGAGTCGAG	1517
Qy	1699	AAATGCAACAGATGAATCCCTTAAGAGAGGTGATTCGAAAATCTGGCTGAGCATATCT	1758
Db	1518	AAACCGCGCAGACGAAGCCTTCCGAAGAGGCTGATCGCAAACTGGCGGAGACATCT	1577
Qy	1759	ATTACATGCTAGCAAGTCCGTGCAATATGTCACACACATTCGTGGCTTGCCCTCTCT	1818
Db	1578	CTTCACCGCAGCAAGTCTCTGCGCTATCATGTCCACCACATTCGTGGCCCTGCTGTCT	1637
Qy	1819	CTACAGACACAGCAGGGAATGATCTCTCCACATTTGGTGGAGACTTCTTTTGTGATGA	1878
Db	1638	CTACAGACACAGCAGGGAATCCACCTCTCCACGCTGTTGGAAGACTTCTTTTGTGATGA	1697
Qy	1879	AGAGGAATCTCTGGCTCGTGAATTTTGACTGGGGTTCCTCAGGAAATTCAGAAATGTAGT	1938
Db	1698	GGAGGAATCTCTAGCTCGGATTTTGACTCTGGGCTCTCCGGGAAATTCAGAAATGTAGT	1757
Qy	1939	AAATGATGCCATACAGCTGCTGGGAAATTTGTTCACAATCACCACACTAGCAGGACGA	1998
Db	1758	CATGCATGCTATTTCAGCTTCTGGGGAATCTGTGTACAATCACCACACTAGCAGGAGGA	1817
Qy	1999	TGAGTTTTTTTATCACCCCCAGCAAACTGTGCCATCATGTTCCGAATCTCAACTTCTACAG	2058
Db	1818	TGAATTTCTTTTACTCCAGCAAACTGTCCCGTCTCTTGAACTCAACTTCTACAG	1877
Qy	2059	CAATGGGGTACTTCATGCTCTTTATCATGGAAGGCCATCATAGTTCGAGCCTTTATGCAGT	2118
Db	1878	CAATGGGGTACTTCATGCTCTTTATCATGGAAGGCCATCATAGTTCGAGCCTTTATGCAGT	1937

Qy	2119	TCTGAA	CAAGAGGGG	ACTGGGGGGTCC	CACTAGCA	CCCA	CACTAA	CCTGAT	CAGCAGGA	2178	
Db	1938	CCAGAA	TAAGAGGGG	TTCCGGAAGGG	CTGCGGAGGC	CTTGG	CAATCT	GATCAGC	CAGCAGGA	1997	
Qy	2179	GCAGCT	TGTCGGAAG	CGGCCCA	CGCTGTGCTAT	CTTCTCT	CAATGA	AGGACCA	CTCTC	2238	
Db	1998	GCAGCT	TGTCGGAAG	CGGCCCA	CGCTGTGCTAT	CTTCTCT	TAATGA	AGGATACCA	ATTTTC	2057	
Qy	2239	ACTGCT	TGTCGACACAT	TTTTACCA	AGTCTCTGCCAT	GAAACAG	TAGGAA	AGTTTAT	TCCAGTA	2298	
Db	2058	TCTGCC	CTGCCAGACAT	TTTACCA	GCTTGTCA	AGAGACA	GATAGGAA	AGTTCA	TCCAGTA	2117	
Qy	2299	TGGCAT	TTCTTACAGT	TGCGAGAGCA	CGATGAC	CAAGGA	AGATATC	AGTCTTAG	TTTCTCTGA	2358	
Db	2118	CGGAAT	TTCTCACAGT	TGCGAGAGCA	AGATGAC	CAAGGA	AGATGT	CAGTCTCG	CTTGCAGA	2177	
Qy	2359	GCAGCAGT	TGGGACAA	GAAGCTTCCAGAA	CCTTTGTCTT	TGGAGNA	AGTGATG	AGAAATGA		2418	
Db	2178	GCAGCAGT	TGGAACA	GAAGCTTCCGAGGC	CTTGAACT	CGGAGAA	GTGACGA	AGAAATGA		2237	
Qy	2419	AGACAGT	GACTTTTGGGAG	GAAACAG	CGAGATTC	TACTGAA	GTGCAG	CAATCC	CAAGGA	2478	
Db	2238	GGA	CAGTGA	CTTTGTGAGAG	CGAGCGTGAT	TCTACTGA	AGTGAGC	CGACG	CCCAAGGA	2297	
Qy	2479	GCACCAG	CAGTTTAT	CACCTTCTTAC	AGAGACTC	CTTGGG	CCCTTCTGCT	GGAGCC	TACAG	2538	
Db	2298	GCACCAG	CAATTAT	CACTCTTCTG	CAGAGGCTTCT	TGGG	CCCTGCTAG	AGACCT	TACAG	2357	
Qy	2539	CTCTGT	TGCCATCTTT	TTTCAACAT	TTCA	AGTGTCTCT	TGTCAGAA	CCCTGAGT	ATCTGCA	2598	
Db	2358	CTCTGT	TGCCATCTTT	TTTCCACAC	CTTCCG	CGCCCA	GTC	CGGAGTCT	GAGTACCTGCA	2417	
Qy	2599	AAAGTT	TGCAAA	ATACCTAAT	ACAGAA	CAGAA	AAATGTT	TGAGTAT	TGCTCAGAG	2658	
Db	2418	GAAGCT	TGCAAGT	ATACCTTCTCA	CCAGGAC	GGAGAGAA	CGTCC	CGGCTGTA	CGCTGAGAG	2477	
Qy	2659	TGCCACA	TATGTC	TTGTGAAGAA	TGCTGTGAAA	TGTTTAA	GGAATAT	TGGGGT	TTTTCAA	2718	
Db	2478	TGCCA	CATACTGCTT	TGTGNA	GAATGCTGTGAAA	ATGTTTAA	GGACAT	TCGGGG	TTTTTCAA	2537	
Qy	2719	GGAGAC	CAAAAC	TAAGAGAGT	GTCTCTTTT	TAGAACT	GTAGCAG	CAC	TTTTCTACCT	CAATG	2778
Db	2538	AGAGAC	CAAGCA	AGCGAG	CGCTCTGCTT	TAGAACT	GTAGCA	CCACTTT	CTCTACCT	CAGG	2597
Qy	2779	CAACCGA	CAAAA	CTTCTT	AGATAATAT	TTCTG	AGTTTTT	TGCTG	CTGTAG	2827	
Db	2598	CAGCCG	CGAAG	CGCTCT	TGAAATAC	ATCTG	AGCTTC	TGTTGGT	TGCTGTAG	2646	

RESULT 9

RESULI 5  
ABA08415

ABA08413  
ID : ABA08415 standard: cDNA: 469 bp.

XX  
CT  
ABAC06413

AC ABA08415;

XX  
XX

DT 11-JAN-2002 (first entry)

2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841

DE Human sn-glycerol-3-P acyltransferase homologue cDNA, SEQ ID NO:191.

Human; cytokine; cell proliferation; cell differentiation; growth factor;  
 haematopoiesis regulation; tissue growth; immunomodulator; activin;  
 inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
 proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
 myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 chronic inflammatory condition; proliferative retinopathy;  
 atherosclerosis; coronary heart disease; arterial ischaemia;  
 bone disorder; osteoporosis; vascular growth disorder;  
 tissue regeneration; wound healing; infection; immune disorder;  
 cell culture; drug screening; gene therapy; antiinflammatory;  
 cytostatic; osteopathic; haemostatic; antiarteriosclerotic;  
 antifungal; vulnery; antitumor; ss.





CC Also included are; (1) a transformed cell having a nucleic acid  
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-  
CC translated sequence that functions in the cell to cause termination of  
CC transcription and addition of polyadenylated ribonucleotides to a 3' end  
CC of the mRNA molecule; and (2) determining a level or pattern of a  
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker  
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its  
CC complement or fragment) with a complementary nucleic acid molecule  
CC obtained from the bovine cell or tissue, where hybridisation between the  
CC marker nucleic acid and the complementary nucleic acid permits the  
CC detection of the molecule; and (b) detecting the level or pattern of the  
CC complementary nucleic acid, where the detection of the complementary  
CC nucleic acid is predictive of the level or pattern of the molecule.  
CC The LMFD nucleic acid is used for determining a level or pattern  
CC of a molecule in a bovine cell or tissue. It is useful for genome  
CC mapping, gene identification and analysis, cattle breeding, preparation  
CC of constructs for use in cattle gene expression, or for genetically  
CC improving cattle. The present sequence is one of the 15112 bovine  
CC LMFD EST (expressed sequence tag) nucleic acids.  
CC Note: The present sequence was not shown in the specification but  
CC was obtained in electronic format from the USPTO web site:  
CC seqdata.uspto.gov/sequence.html?DocID=20020137139.  
XX  
SQ Sequence 429 BP; 113 A; 106 C; 94 G; 116 T; 0 other;

Query Match 11.4%; Score 342.2; DB 25; Length 429;  
Best Local Similarity 90.2%; Pred. No. 2e-95;  
Matches 378; Conservative 0; Mismatches 38; Indels 3; Gaps 1;  
QY 1670 GGTAGACACAGTCCATTAAAGTCCAGAAATGCAACAGATGAATCCTAGAGGAGG 1729  
DB 14 GATACAGACAGTCCATTAAAGTCCAGAAATGCAACAGATGAATCCTAGAGGAGG 70  
QY 1730 TTGATTGCAAAATCTGGCTGAGCATATTTCTATTCACTGTAGCAAGTCTGTGCCATTATG 1789  
DB 71 CTGATTGCACATCTGGCTGAGCACATTTCTTCACTGTAGCAAGTCTGTGCCATTATG 130  
QY 1790 TCCACACACATTTGGTGTGCTGCTCTCTACAGACACAGGAGGAATTTGATCTCTCC 1849  
DB 131 TCCAGCACATCGTGACCTGTCTGCTCTCTACAGACACAGGAGGAATTTGCTCTCC 190  
QY 1850 ACATTTGTCGAGACTTTCTTTGTGATGAAGAGGAAGTCTGGCTGTGATTTGACCTG 1909  
DB 191 ACATTTGTCGAGACTTTCTTTGTGATGAAGAGGAAGTCTGGCTGTGATTTGACCTG 250  
QY 1910 GGGTTCTCAGGAAATTCAGAGATCTAGTAATGATGATGATGATGATGATGATGATGAT 1969  
DB 251 GGTCTCTCAGGAAATTCAGAGATCTAGTAATGATGATGATGATGATGATGATGATGAT 310  
QY 1970 GTCAATCACCACACTAGCAGGAAATGATGATGATGATGATGATGATGATGATGATGATG 2029  
DB 311 ATCAATCACCACACTAGCAGGAAATGATGATGATGATGATGATGATGATGATGATGATG 370  
QY 2030 CCATCAGTCTTGAACACTAACTTCTACAGCAATGGGGTACTTCTTATCATGGA 2088  
DB 371 GCATCAGTCTTGAACACTAACTTCTACAGCAATGGGGTACTTCTTATCATGGA 429

## RESULT 11

ABX43656  
ID ABX43656 standard; cDNA; 407 BP.  
XX  
XX AC ABX43656;  
XX  
XX 21-FEB-2003 (first entry)  
XX  
DE Bovine EST associated with lactation/muscle/fat deposition #8821.  
XX  
XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
KW muscle deposition; fat deposition; genome mapping; gene identification;  
KW gene analysis; cattle breeding.  
XX  
OS Bos Taurus.

XX US2002137139-A1.

XX 26-SEP-2002.

XX 24-SEP-2001; 2001US-0960352.

XX 12-JAN-1999; 99US-115707P.

XX 11-JAN-2000; 2000US-0480902.

XX (BYAT/) BYATT J C.

XX (MATH/) MATHIALAGAN N.

XX (TAON/) TAO N.

XX (WARR/) WARREN W C.

XX Byatt JC, Mathialagan N, Tao N, Warren WC;

XX WPI; 2003-110599/10.

XX New nucleic acid associated with lactation, and muscle and fat  
XX deposition, useful for genome mapping, gene identification and  
XX analysis, cattle breeding, or for genetically improving cattle

XX Claim 2; SEQ ID No 8821; 245pp; English.

XX The invention relates to a purified nucleic acid molecule associated with  
XX lactation or muscle and fat deposition (designated LMFD), derived  
XX from cattle, and the LMFD nucleic acid can specifically hybridise to a  
XX second nucleic acid molecule comprising any of 15112 nucleotide  
XX sequences, appearing as ABX34836-ABX49947, or complements of them.  
XX Also included are; (1) a transformed cell having a nucleic acid  
XX comprising an LMFD nucleic acid linked to a promoter and a 3' non-  
XX translated sequence that functions in the cell to cause termination of  
XX transcription and addition of polyadenylated ribonucleotides to a 3' end  
XX of the mRNA molecule; and (2) determining a level or pattern of a  
XX nucleic acid (comprising any of the 15112 nucleic acid sequences or its  
XX complement or fragment) with a complementary nucleic acid molecule  
XX obtained from the bovine cell or tissue, where hybridisation between the  
XX marker nucleic acid and the complementary nucleic acid permits the  
XX detection of the molecule; and (b) detecting the level or pattern of the  
XX complementary nucleic acid, where the detection of the complementary  
XX nucleic acid is predictive of the level or pattern of the molecule.  
XX The LMFD nucleic acid is used for determining a level or pattern  
XX of a molecule in a bovine cell or tissue. It is useful for genome  
XX mapping, gene identification and analysis, cattle breeding, preparation  
XX of constructs for use in cattle gene expression, or for genetically  
XX improving cattle. The present sequence is one of the 15112 bovine  
XX LMFD EST (expressed sequence tag) nucleic acids.  
XX Note: The present sequence was not shown in the specification but  
XX was obtained in electronic format from the USPTO web site:  
XX seqdata.uspto.gov/sequence.html?DocID=20020137139.

XX Sequence 407 BP; 108 A; 101 C; 89 G; 109 T; 0 other;

Query Match 11.2%; Score 337.6; DB 25; Length 407;

Best Local Similarity 90.9%; Pred. No. 5.2e-94;

Matches 371; Conservative 0; Mismatches 34; Indels 3; Gaps 1;

QY 1670 GGTAGACACAGTCCATTAAAGTCCAGAAATGCAACAGATGAATCCTAGAGGAGG 1729

DB 3 GATACAGACAGTCCATTAAAGTCCAGAAATGCAACAGATGAATCCTAGAGGAGG 59

QY 1730 TTGATTGCAAAATCTGGCTGAGCATATTTCTATTCACTGTAGCAAGTCTGTGCCATTATG 1789

DB 60 CTGATTGCACATCTGGCTGAGCACATTTCTTCTACTGTAGCAAGTCTGTGCTATTATG 119

QY 1790 TCCACACACATTTGGTGTGCTGCTCTCTCTACAGACACAGGAGGAATTTGATCTCTCC 1849

DB 120 TCCAGGACATCTGGGCTGTCTCTCTCTCTACAGACACAGGAGGAATTTGCTCTCTCC 179

QY 1850 ACATTGTCGAAGACTTTTGTGATGAAAGAGGAAGTCTGTGCTGTGATTTTGCACCTG 1909

Db 180 ACATTGGTGAAGACTTCTTTGTGATGAAGAGGAAAGTCTCGTCTGATTTTGACTTG 239  
Qy 1910 GGGTTCTCAGAAATTCAGAAAGATGTAGTAATGATGCCATACAGCTGCTGGGAAATTGT 1969  
Db 240 GGTTCCTCAGAAATTCAGAAAGATGTAGTATGATGCCATACAGCTCTGGGAAATTGT 299  
Qy 1970 GTCAATATCCACACTAGCAGGAACGATGAGTTTTTTATCACCCCGCAGCACAACCTGTC 2029  
Db 300 ATCACAATACCCACAGCAGCAAGATGATGAGTTTTTTTATTACTCTCTAGCACAACCTATC 359  
Qy 2030 CCATCAGTCTTCGAACTCAACTCTCTACAGCAATGGGATCTTCATGTC 2077  
Db 360 CCATCAGTCTTTGAACTCAACTCTCTACAGCAATGGGATCTTCATGTC 407

RESULT 12  
ABX35577/c  
ID ABX35577 standard; cDNA; 418 BP.  
XX AC ABX35577;  
XX AC  
XX 20-FEB-2003 (first entry)  
XX Bovine EST associated with lactation/muscle/fat deposition #742.  
XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
KW muscle deposition; fat deposition; genome mapping; gene identification;  
KW gene analysis; cattle breeding.  
XX Bos Taurus.  
XX US2002137139-A1.  
XX 26-SEP-2002.  
XX 24-SEP-2001; 2001US-0960352.  
XX 12-JAN-1999; 99US-115707P.  
PR 11-JAN-2000; 2000US-0480902.  
XX (BYATT/) BYATT J C.  
PA (MATH/) MATHIALAGAN N.  
PA (TAON/) TAO N.  
PA (WARR/) WARREN W C.  
XX Byatt JC, Mathialagan N, Tao N, Warren WC;  
XX WPI; 2003-110599/10.

XX New nucleic acid associated with lactation, and muscle and fat  
PT deposition, useful for genome mapping, gene identification and  
PT analysis, cattle breeding, or for genetically improving cattle  
XX Claim 2; SEQ ID No 742; 245pb; English.  
XX The invention relates to a purified nucleic acid molecule associated with  
XX lactation or muscle and fat deposition (designated LMFD), derived  
XX from cattle, and the LMFD nucleic acid can specifically hybridize to a  
XX second nucleic acid molecule comprising any of 1512 nucleotide  
XX sequences, appearing as ABX34836-ABX49947, or complements of them.  
XX Also included are: (1) a transformed cell having a nucleic acid  
XX comprising an LMFD nucleic acid linked to a promoter and a 3' non-  
XX translated sequence that functions in the cell to cause termination of  
XX transcription and addition of polyadenylated ribonucleotides to a 3' end  
XX of the mRNA molecule; and (2) determining a level or pattern of a  
XX molecule in a bovine cell or tissue comprising: (a) incubating a marker  
XX nucleic acid (comprising any of the 1512 nucleic acid sequences or its  
XX complement or fragment) with a complementary nucleic acid molecule  
XX obtained from the bovine cell or tissue, where hybridisation between the  
XX marker nucleic acid and the complementary nucleic acid permits the  
XX detection of the molecule; and (b) detecting the level or pattern of the  
XX complementary nucleic acid, where the detection of the complementary  
XX nucleic acid is predictive of the level or pattern of the molecule.

CC The LMFD nucleic acid is used for determining a level or pattern  
CC of a molecule in a bovine cell or tissue. It is useful for genome  
CC mapping, gene identification and analysis, cattle breeding, preparation  
CC of constructs for use in cattle gene expression, or for genetically  
CC improving cattle. The present sequence is one of the 1512 bovine  
CC LMFD EST (expressed sequence tag) nucleic acids.  
CC Note: The present sequence was not shown in the specification but  
CC was obtained in electronic format from the USPTO web site:  
CC seqdata.uspto.gov/sequence.html?DocID=20020137139.

XX SQ Sequence 418 BP; 97 A; 116 C; 102 G; 102 T; 1 other;  
Query Match 10.1%; Score 302; DB 25; Length 418;  
Best Local Similarity 84.1%; Pred. No. 6.4e-83;  
Matches 354; Conservative 0; Mismatches 61; Indels 6; Gaps 1;

Qy 2112 ATGCGATTCTGAACAGAGGGGACTGGGGGTCCCACTAGCACCCCACTACCTGATCA 2171  
Db 418 ATGCGATTCTGAAGAGAGGGGCGCNCAGGAGGACC-----CGCGTCTCCAGCTTGGTCA 365  
Qy 2172 GCCAGGAGCAGCTGGTGGGAAGGCGCCAGCTGTGCTACCTTCTCTCAATGAAGCA 2231  
Db 364 GCCAGGAGCAGCTGGTGGCAAGGCTGCCAGCTGTGCTATCTCTCAATGAAGCA 305  
Qy 2232 CCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTA 2291  
Db 304 CCATCTCTCCCTGCCAGACCTTTTACCAAAATTTGCCATGAACAGTGGCGCGTTTA 245  
Qy 2292 TCCAGTATGGCATTTCTACAGTGGGAGCAGATGACCAGGAATATCAGTCTCTAGTC 2351  
Db 244 TCCAGTATTTGCAATTTTATAGTTGCCAGCAAGATAATCAGGAAGATATCAGTCTCTTGT 185  
Qy 2352 TTGCTGAGCAGCTGGGACAGAAAGCTTCCAGAACCTTTGTCTTGGAGAGTGATGAAG 2411  
Db 184 CTTCGAGCAGCAGTTGGACAAGAGTTCCGGAACCTTTGTCTTGGAGAGTGATGAAG 125  
Qy 2412 AAGATGAACACAGTGAATTTGGGAGGAAACAGCAGATTTGCTACCTGAAGGTGAGCAAT 2471  
Db 124 AAGATGAACACAGCGATTTTGTGAGGAGCAACGAGATTGCTACCTGAAGGTGAGCAAT 65  
Qy 2472 CCAAGGAGCAGCAGTTTATCAGCTTTTACAGAGACTCCTTGGGCTTTGCTGGAGG 2531  
Db 64 CCAAGGAGCAGCAGTTTATCAGCTTTTATCAGAGGCTCCTCGGGCTCTCTGCTGGAGG 5  
Qy 2532 C 2532  
Db 4 C 4

RESULT 13  
AAC05673  
ID AAC05673 standard; cDNA; 271 BP.  
XX AC AAC05673;  
XX AC  
XX 06-OCT-2000 (first entry)  
DT Human secreted protein 5' EST, SEQ ID NO: 9748.  
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX Homo sapiens.  
XX EP1033401-A2.  
XX PD 06-SEP-2000.  
XX 21-FEB-2000; 2000EP-0200610.  
XX 26-FEB-1999; 99US-0122487.  
PR (GEST ) GENSET.  
XX PA

XX Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
XX  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
XX Claim 1; SEQ ID 9748; 71pp + CD-ROM; English.  
XX  
XX The present sequence is one of a large number of 5' ESTs derived from  
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively  
XX identified within the present sequence. The 5' ESTs were prepared from  
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
XX sequences usually correspond mainly to the 3' untranslated region (UTR)  
XX of the mRNA because they are often obtained from oligo-dT primed cDNA  
XX libraries. Such ESTs are not well suited for isolating cDNA sequences  
XX derived from the 5' ends of mRNAs and even in those cases where longer  
XX cDNA sequences have been obtained, the full 5' UTR is rarely included.  
XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX They are used to obtain upstream regulatory sequences and to design  
XX expression and secretion vectors.  
XX  
XX Sequence 271 BP; 69 A; 77 C; 64 G; 61 T; 0 other;  
XX  
XX Query Match 8.2%; Score 246.4; DB 21; Length 271;  
XX Best Local Similarity 99.6%; Pred. No. 9.7e-66;  
XX Matches 247; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX 12 GCTGTCGGGGGGAAGTCTTTCTGAGGTTACTGTGGAGCACCACCAAGTCTGTGACCTCT 71  
XX  
XX 24 GCTGTCGGGGGGAAGTCTTTCTGAGGTTACTGTGGAGCACCACCAAGTCTGTGACCTCT 83  
XX  
XX 72 GGCCCTGCAACAGAGCACCAGAGAAACAGACCTTGTATTATCCACAGCCTGGGAC 131  
XX  
XX 84 GGCCCTGCAACAGAGCACCAGAGAAACAGACCTTGTATTATCCACAGCCTGGGAC 143  
XX  
XX 132 TGCTCTCTCCAGAGTCTCCATCAGCTTTCCTTAATGACTGATGGAAATATCTCTCAA 191  
XX  
XX 144 TGCTCTCTCCAGAGTCTCCATCAGCTTTCCTTAATGACTGATGGAAATATCTCTCAA 203  
XX  
XX 192 CACCACCAAGTCAAGGATACAGGACAGCGGCTCCCTGTTGTATGACATTTCTGACC 251  
XX  
XX 204 CACCACCAAGTCAAGGATACAGGACAGCGGCTCCCTGTTGTATGACATTTCTGACC 263  
XX  
XX 252 CGAAACTG 259  
XX  
XX 264 CGAAACCG 271  
XX  
XX  
XX RESULT 14  
XX ID ABX42469 standard; cDNA; 387 BP.  
XX  
XX AC ABX42469;  
XX  
XX  
XX 20-FEB-2003 (first entry)  
XX  
XX DE Bovine EST associated with lactation/muscle/fat deposition #7634.  
XX  
XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
XX muscle deposition; fat deposition; genome mapping; gene identification;  
XX gene analysis; cattle breeding.  
XX  
XX OS Bos Taurus.  
XX  
XX PN US2002137139-A1.  
XX  
XX 26-SEP-2002.  
XX  
XX

PF 24-SEP-2001; 2001US-0960352.  
XX  
XX 12-JAN-1999; 99US-115707P.  
XX 11-JAN-2000; 2000US-0480902.  
XX  
XX (BYAT/) BYATT J C.  
XX (MATH/) MATHIALAGAN N.  
XX (TAON/) TAO N.  
XX (WARR/) WARREN W C.  
XX  
XX Byatt JC, Mathialagan N, Tao N, Warren WC;  
XX WPI; 2003-110599/10.  
XX  
XX New nucleic acid associated with lactation, and muscle and fat  
XX deposition, useful for genome mapping, gene identification and  
XX analysis, cattle breeding, or for genetically improving cattle -  
XX  
XX Claim 2; SEQ ID No 7634; 245pp; English.  
XX  
XX The invention relates to a purified nucleic acid molecule associated with  
XX lactation or muscle and fat deposition (designated LMFD), derived  
XX from cattle, and the LMFD nucleic acid can specifically hybridise to a  
XX second nucleic acid molecule comprising any of 15112 nucleotide  
XX sequences, appearing as ABX34836-ABX49947, or complements of them.  
XX Also included are; (1) a transformed cell having a nucleic acid  
XX comprising an LMFD nucleic acid linked to a promoter and a 3' non-  
XX translated sequence that functions in the cell to cause termination of  
XX transcription and addition of polyadenylated ribonucleotides to a 3' end  
XX of the mRNA molecule; and (2) determining a level or pattern of a  
XX nucleic acid in a bovine cell or tissue comprising: (a) incubating a marker  
XX nucleic acid (comprising any of the 15112 nucleic acid sequences or its  
XX complement or fragment) with a complementary nucleic acid molecule  
XX obtained from the bovine cell or tissue, where hybridisation between the  
XX marker nucleic acid and the complementary nucleic acid permits the  
XX detection of the molecule; and (b) detecting the level or pattern of the  
XX complementary nucleic acid, where the detection of the complementary  
XX nucleic acid is predictive of the level or pattern of the molecule.  
XX The LMFD nucleic acid is used for determining a level or pattern  
XX of a molecule in a bovine cell or tissue. It is useful for genome  
XX mapping, gene identification and analysis, cattle breeding, preparation  
XX of constructs for use in cattle gene expression, or for genetically  
XX improving cattle. The present sequence is one of the 15112 bovine  
XX LMFD EST (expressed sequence tag) nucleic acids.  
XX Note: The present sequence was not shown in the specification but  
XX was obtained in electronic format from the USPTO web site:  
XX seqdata.uspto.gov/sequence.html?docID=20020137139.  
XX  
XX SQ Sequence 387 BP; 87 A; 97 C; 113 G; 90 T; 0 other;  
XX  
XX Query Match 8.2%; Score 244.8; DB 25; Length 387;  
XX Best Local Similarity 79.6%; Pred. No. 3.9e-65;  
XX Matches 301; Conservative 0; Mismatches 74; Indels 3; Gaps 1;  
XX  
XX 2614 CCTAATAACAGAGAACAGAGAAATGTTGCAGTATATGCTGAGAGTGCCATATTGTCT 2673  
XX  
XX 1 CCTGATCAGAGAACCCGAGAGAGCGGTTCGCTGCTACGCTGAGAGCGCCATTACTGTCT 60  
XX  
XX 2674 TGTGAAGATGCTGTGAAATGTTTAAGGATATTTGGGGTTTTCAAGGAGACCAACAAA 2733  
XX  
XX 61 CGTGAAGATGCTGTGAAAGACGTTTAAGGATATCGGGGTTTTCAAGGAGACCAACAAA 120  
XX  
XX 2734 GAGAGTCTGTGTTTTAGAACCTGACGACACCTTTTCTACCTCAATGACACCAAAACT 2793  
XX  
XX 121 GAGAGTGTGTGTTTTAGAACCTGACGACACCTTTTCTACCTCAATGACACCAAAACT 180  
XX  
XX 2794 TCTAGAATATATTCTGAGTGTGTTTGGTGTCTAGTAACTGCTGCTGCAATG 2853  
XX  
XX 181 CCTGAATATATTCTGAGTGTGTTTGGTGTCTAGTAACTGCTGCGGACCTCTGCAAGTG 240  
XX  
XX 2854 AAGGTCTATGATGAGTTCCTTGTAGGTACAGCTTCTGGCTCAAGAGTTTGAAGGTGCC 2913  
XX  
XX 241 AAGGAGCGGACCGGTTCTCTCGTAGG--CCCGCGCTGTGGCAGAGAGTGGACGGCGC 297  
XX  
XX



RESULT 2  
US-09-252-991A-4533/c  
; Sequence 4533, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18





Db 807868 GTGTAATTAAAAAATACGTAA 807847

RESULT 5

US-09-643-990A-1/c

; Sequence 1, Application US/09643990A

; Patent No. 6528289

GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann,

; Mark D. Adams

; Owen White

; Hamilton O. Smith

; J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of

; the Haemophilus influenzae Rd Genome, Fragments

; Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville,

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/643,990A

FILING DATE: 23-Aug-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,429

FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787

FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB186P1C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-610-5790

TELEFAX: 310-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-643-990A-1

Query Match 1.5%; Score 44.8; DB 4; Length 1830121;

Best Local Similarity 45.2%; Pred. No. 0.19;

Matches 227; Conservative 0; Mismatches 257; Indels 18; Gaps 1;

Qy 1011 TCTGTTTCTACCAAGTTTCATAGATCCCATATGATCTGCTGCTCACTTTCTCTCT 1070

Db 808330 TTGTTTATGTGCTTCCGCCACCGTAGCTATGCTATGCTATGCTTCTTCTTATGCTTT 808271

Qy 1071 TCTGCCATAACATCAAGCACCATACATGTTCTCAGGCAATAATCTCAACATCCCAATCT 1130

Db 808270 ATCATCAAGGCTTGTTCGCCACATATCGCGGAGGATTAACCTTAATTTCTGCGCTTA 808211

Qy 1131 TCAGTACCTTGATCCATAGCTTTGGGGGCTTCTTCATACGACGAAGGCTCGATGAACAC 1190

Db 808210 TAGGTAGATGTTTCGTAGTTGGGGGCTTCTTTATTCGCCGACCTTTTAAGGGAATC 808151

Qy 1191 CAGATGGACGGAAGATGTTCTCTATAGAGCTTCTCTATAGAGCTTCTCCTATGGGCAATAGTGAATPAC 1250

Db 808150 GA-----CTATATCTGCAATTTTTCGCCGAATATTTATCAGAATCAT 808109

Qy 1251 TTCACAGACGCAATTTCTTGAGATCTTCTCGAAGGCACACGTTCTAGGAGTGAAGAAA 1310

Db 808108 TCCATCGAGGCTATTTCGTCGAATATTTATTAGGGGGGCTGTTCTCGTACAGGTCGTT 808049

Qy 1311 CCTCTTGCTCGGGCAGGACTTTTTCAGTTGTTGTTAGATACCTCTGTCTACCAATGTCA 1370

Db 808048 TGCTGGCACCAAAACACAGGTATGATGTCAATGACACTTCAAGCATTTGCAACATAGTCAA 807989

Qy 1371 TCCCAGACATCTTGATTAATACCTGTGGAAATCTCTATGATCGCATTTACGAAGTCACT 1430

Db 807988 CTGCCCCCAATTCGATTTGTTCTGTTTACGTTGTTTGAACACGTATTAGAAGTAGACA 807929

Qy 1431 ACAATGTTGAACACTGGGCAACCTAAGAAAGAAATGAGAGCTGTGGAGTGTAGCAAGAG 1490

Db 807928 CTTAATGCTAAGGAATTACGTGTCAGCGAAAGAAAGAAATCGCGGTTTAGTACTTC 807869

Qy 1491 GTGTTATTAGAAATGTTAGAAA 1512

Db 807868 GTGTAATTAAAAAATACGTAA 807847

RESULT 6

US-08-916-421B-1/c

; Sequence 1, Application US/08916421B

; Patent No. 6503729

GENERAL INFORMATION:

APPLICANT: Bult et al.

TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus

; Patent No. 6503729

TITLE OF INVENTION: jannaschii

FILE REFERENCE: PB275

CURRENT APPLICATION NUMBER: US/08/916,421B

CURRENT FILING DATE: 1997-08-22

PRIOR APPLICATION NUMBER: US 60/024,428

PRIOR FILING DATE: 1996-08-22

NUMBER OF SEQ ID NOS: 3

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 1664976

TYPE: DNA

ORGANISM: Methanococcus jannaschii

FEATURE:

NAME/KEY: misc feature

LOCATION: (28222)..(28222)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (28257)..(28258)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (84773)..(84773)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (84808)..(84808)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (84812)..(84812)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (98120)..(98120)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (98159)..(98159)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (98239)..(98239)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (98266)..(98266)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (98343)..(98343)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature  
LOCATION: (103998)..(103998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (148948)..(148948)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (163385)..(163385)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191989)..(191989)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191995)..(191995)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (231980)..(231980)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234220)..(234220)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature

LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1470091)..(1470091)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1664855)..(1664855)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-421B-1

Query Match 1.3%; Score 38.6; DB 4; Length 1664976;

Best Local Similarity 50.8%; Pred. No. 18;

Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 678 TTTTATTCAAGAGCGAGATGTGCATAAGGGCATGTTGCCACCAATGTGACTGAAATG 737  
|||||  
Db 541143 TCTTTATTAAACATAGAGAGGATAACAGATGCATGATATAAGTAGTGAATAGTCAATG 541084  
QY 738 TCGTGAACAGCAGTAGAGTACAAAGGCGCAATGGCAGAGTGGCTGCTGAATTAACCCCTG 797  
|||||  
Db 541083 CTCTTAAATGAAATATTAAAGAAATAACAAAATTTGTCAGTTGTAGGAGAGCCAA 541024  
QY 798 ATGGTTCTGCCAGCAGCAATCAAAAGCCGTTAAACAAAGTGAAGAAAGCTTAAAGGA 857  
|||||  
Db 541023 AAACCTCTGAAAGTTGGATACAATAATGTTGGTTAATAAGAGCATGTGAACACTA 540964  
QY 858 T 858

Db 540963 T 540963

RESULT 7

US-09-328-352-2557

Sequence 2557, Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 2557

LENGTH: 1578

TYPE: DNA

ORGANISM: Acinetobacter baumannii

US-09-328-352-2557

Query Match 1.2%; Score 37.2; DB 4; Length 1578;

Best Local Similarity 54.3%; Pred. No. 0.32;

Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 728 ACTGAAATGTGCTGAACAGCAGTAGAGTACAGAGGCAATTCGACAGAGTGGCTGCTGAA 787

Db 1189 ATTGAAATGTACCGAAGCAGTCTTCTATGCTAGGCGCAATGTCTTGGTTTATGACTCA 1248

Qy 788 TTAACCTTGATGTTCTGCCAGCAGCAATCAAAAGCGGTTAAACAAAGTGAAGAAA 847

Db 1249 TTCATGTTCTTGTGGAAGCCCACTGCAATATATTGTAGATGTTTATTGCAATGAAA 1308

Qy 848 GCTAAAGGATTTCTCAA 865

Db 1309 GCGAAGGTATTGCTAAA 1326

RESULT 8

US-08-895-601-2

Sequence 2, Application US/08895601

Patent No. 6060262

GENERAL INFORMATION:

APPLICANT: Beer-Romero, Peggy

APPLICANT: Strack, Peter J.

APPLICANT: Glass, Susan J.

APPLICANT: Rolfe, Mark

TITLE OF INVENTION: REGULATION OF KAPPA B (IKB) DEGRADATION,

TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/895,601

FILING DATE: 16-JUL-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: MIV-096.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-832-1000

TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2790 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 2...2782

US-08-895-601-2

Query Match 1.2%; Score 36.6; DB 3; Length 2790;

Best Local Similarity 55.0%; Pred. No. 0.77;

Matches 72; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 735 ATGTGCTGAACAGCAGTAGAGTACAGAGGCAATTCGACAGAGTGGCTGCTGAATTAACC 794

Db 531 ATTTACCTAAACCAAGTGGCTCAGAAGATGATAATGCAAGAACAGGCTGAGGAATTAGAGC 590

Qy 795 CTGATGTTCTGCCCCAGCAGCAATCAAAAGCGGTTAAACAAAGTGAAGAAAAGCTAAA 854

Db 591 CTGGCTGGTGTGTTTGGACCAACCAAGATGCTGTGTCATTTCAGCAACCAAGAAC 650

Qy 855 GGATTCTTCAA 865

Db 651 CTTCTCCTCTA 661

RESULT 9

US-08-232-463-14/c

Sequence 14, Application US/08232463

Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.

APPLICANT: SCHEIFLINGER, F.

APPLICANT: FALKNER, F. G.

TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,463

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,313

FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/114 IMMU

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: ptzgt-F18  
US-08-232-463-14

Query Match 1.2%; Score 35.6; DB 1; Length 7218;  
Best Local Similarity 1.5%; Pred. No. 3.3;  
Matches 5; Conservative 185; Mismatches 134; Indels 0; Gaps 0;

QY 679 TTTTATCAAGAGGAGATGTGATAGGCGCATGTTGCCACCAATGTGACTGAAATGT 738  
DB 1442 TTTGGTACRR 1383  
QY 739 GCTGAACAGCAGTAGAGTCAAGAGGCAATTCAGAGAGTGGCTGCTGAAATTAACCCCTGA 798  
DB 1382 RRR 1323  
QY 799 TGGTTCTGCCAGCAGCAATCAAAAGCGGTTAAACAAAGTGAAGAAAGCTAAAGGAT 858  
DB 1322 RRR 1263  
QY 859 TCTTCAAGAAATGGTTCCTACTGCTCACCGCAATGATCAGACTGACTGGTGGTGCT 918  
DB 1262 RRR 1203  
QY 919 GCTAAACTGTTCAACAGCTTCTTTTGGACATTTCAAAATTCACAAAGGTCAACTTGAGAT 978  
DB 1202 RRR 1143  
QY 979 GGTTAAAGCTGCAACTGAGACGAA 1002  
DB 1142 RRR 1119

## RESULT 10

US-09-328-352-2440/c  
Sequence 2440, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 2440  
LENGTH: 1032  
TYPE: DNA  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-2440

Query Match 1.2%; Score 35.2; DB 4; Length 1032;  
Best Local Similarity 48.5%; Pred. No. 1.1;  
Matches 97; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 682 TATTCAAGCGAGATGTGATAGGCGCATGTTGGCCACCAATGTGACTGAAATGTGCT 741  
DB 425 TATTCCAGCTCTTTAATTGTCATCATTTGAATCATTTATGCAATTTTGAATGGATGGCTTCT 366  
QY 742 GAACAGCAGTAGAGTCAAGAGGCAATTCAGAGAGTGGCTGCTGAAATTAACCCCTGATGG 801  
DB 365 TGACCAATCAGCGATGTAACACGCTGTTTAAATCGATCTTGAATAAATTCACGGTGG 306  
QY 802 TTCTGCCAGCAGCAATCAAAAGCCGTTTAAACAAAGTGAAGAAAGCTAAAGGATTCCT 861  
DB 305 TGACGAGCTGTTCTATCACCAGATCTTCACCATAGGTCAAAACCCGAAAGCGTTGCA 246  
QY 862 TCAGAAATGGTTCCTACTG 881  
DB 245 AAGAACAAAGTTGCAACTG 226

## RESULT 11

US-09-734-674-3  
Sequence 3, Application US/09734674  
Patent No. 6498022  
GENERAL INFORMATION:  
APPLICANT: WEI, Ming-Hui et al  
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
TITLE OF INVENTION: AND USES THEREOF  
FILE REFERENCE: CL001018  
CURRENT APPLICATION NUMBER: US/09/734,674  
CURRENT FILING DATE: 2000-12-13  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 202001  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)-(202001)  
OTHER INFORMATION: n = A,T,C or G  
US-09-734-674-3

Query Match 1.2%; Score 34.8; DB 4; Length 202001;  
Best Local Similarity 60.6%; Pred. No. 69;  
Matches 57; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 2556 TTCACAACCTTCTAGTGGTCTCTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACC 2615  
DB 32955 TTCTCTAGTCTTTAGGCCAGTGTATAATGTTAGGATTTACAAAAGTTGGTAATATAGA 33014  
QY 2616 TAATACCAAGACAGAAAGAAATGTTGCAGTATA 2649  
DB 33015 GAGAAACAGGAGAAATGAAATGGACAGGAAA 33048

## RESULT 12

US-09-328-352-3343/c  
Sequence 3343, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 3343  
LENGTH: 825  
TYPE: DNA  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-3343

Query Match 1.1%; Score 34.2; DB 4; Length 825;  
Best Local Similarity 52.4%; Pred. No. 1.9;  
Matches 75; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 1688 AATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGAGGTGATTCGAAATCTGGCT 1747  
DB 364 ATTGATGTAACAAAGCGAGACTTGTATAACCTCGCAATAGTTGCTTCTTGGGGCAT 305  
QY 1748 GAGCATTTCTTATCTACTGCTAGCAAGTCTGTGCCATTATGTCACACACATTTGGCT 1807  
DB 304 TACTCACTTCTATATACAAACTATAAGTACTGCTTTAGATAACCATCTGATGTGTT 245  
QY 1808 TGCCTGCTCTCTACAGACACAG 1830  
DB 244 TGCCTGCCGTATCCATACCG 222

## RESULT 13

US-08-749-903-2  
; Sequence 2, Application US/08749903  
; Patent No. 5759812  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hawkins, Phillip R.  
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/749,903  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0163 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1711 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 5759812e  
; IMMEDIATE SOURCE:  
; CLONE: 989953  
US-08-749-903-2  
Query Match 1.1%; Score 34; DB 1; Length 1711;  
Best Local Similarity 63.4%; Pred. No. 3.8;  
Matches 52; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
Qy 1598 GCTCTACTTTCCCTGGAGCAAGCGTTGTACAGCTATCTTCAAGACCCAGTGAT 1657  
Db 1270 GCTGTACAGTGCCTGGGACAGCAGTTTACCCCTGATCTCATCAGGGAAGGCTCTGTGAT 1329  
Qy 1658 GCTGCTGATGAAGGTAGAGACA 1679  
Db 1330 GCTGCAGGTTGATGTAGACACA 1351  
RESULT 14  
US-08-641-2  
; Sequence 2, Application US/09088641  
; Patent No. 6312895  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hawkins, Phillip R.  
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304

US-09-903-2  
; Sequence 2, Application US/08749903  
; Patent No. 5759812  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hawkins, Phillip R.  
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/749,903  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0163 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1711 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 6312895e  
; IMMEDIATE SOURCE:  
; CLONE: 989953  
US-09-088-641-2  
Query Match 1.1%; Score 34; DB 4; Length 1711;  
Best Local Similarity 63.4%; Pred. No. 3.8;  
Matches 52; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
Qy 1598 GCTCTACTTTCCCTGGAGCAAGCGTTGTACAGCTATCTTCAAGACCCAGTGAT 1657  
Db 1270 GCTGTACAGTGCCTGGGACAGCAGTTTACCCCTGATCTCATCAGGGAAGGCTCTGTGAT 1329  
Qy 1658 GCTGCTGATGAAGGTAGAGACA 1679  
Db 1330 GCTGCAGGTTGATGTAGACACA 1351  
RESULT 15  
US-08-536-1/c  
; Sequence 1, Application US/08227536  
; Patent No. 5658784  
; GENERAL INFORMATION:  
; APPLICANT: Eckner, Richard  
; APPLICANT: Ewen, Mark  
; APPLICANT: Livingston, David  
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION  
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
; STREET: Ten Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/227,536  
; FILING DATE: 14-APR-1994  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Ph.D., Kathleen A.  
; REGISTRATION NUMBER: 34,380  
; REFERENCE/DOCKET NUMBER: DPCI-308XX





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2004, 04:06:03 ; Search time 695.063 Seconds  
(without alignments)  
15003.329 Million cell updates/sec

Title: US-09-935-290-1  
Perfect score: 3003  
Sequence: 1 ttccggaccaggctgctgcy.....tctgagatcttccagagttt 3003

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2276164 seqs, 1736306516 residues

Total number of hits satisfying chosen parameters: 4552328

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3001	99.9	3003	11	US-09-935-290-1
2	2970.2	98.9	4033	13	US-09-814-353-20793
3	2958	98.5	3210	13	US-10-288-252-26
4	2484	82.7	2484	11	US-09-935-290-3
5	1896.2	63.1	2646	10	US-09-917-800A-1648
6	581.6	19.4	612	10	US-09-833-381-1961
7	480	16.0	868	10	US-09-833-381-1958
8	437	14.6	482	11	US-09-918-995-21644
9	385.4	12.8	399	11	US-09-918-995-7822
10	348.4	11.6	440	10	US-09-833-381-1964
11	346.6	11.5	396	10	US-09-833-381-1959
12	342.2	11.4	429	10	US-09-960-352-7384
13	337.6	11.2	407	10	US-09-960-352-8821
14	302	10.1	418	10	US-09-960-352-742
15	268.6	8.9	379	13	US-09-814-353-16182

16	244.8	8.2	387	10	US-09-960-352-7634	Sequence 7634, Ap
17	233	7.8	303	10	US-09-960-352-13086	Sequence 13086, A
18	209.6	7.0	264	10	US-09-960-352-6239	Sequence 6239, Ap
19	207.2	6.9	244	13	US-09-814-353-3482	Sequence 3482, Ap
20	207.2	6.9	244	13	US-09-814-353-9798	Sequence 9798, Ap
21	171	5.7	501	10	US-09-960-352-4061	Sequence 4061, Ap
22	129	4.3	308	10	US-09-833-381-1963	Sequence 1963, Ap
23	109	3.6	2682	13	US-10-353-690-31	Sequence 31, Appl
24	109	3.6	2755	13	US-10-288-252-27	Sequence 27, Appl
25	93.6	3.1	399	10	US-09-960-352-54	Sequence 54, Appl
26	83.2	2.8	404	10	US-09-960-352-5756	Sequence 5756, Ap
27	75.2	2.5	467	13	US-10-027-632-297935	Sequence 297935, A
28	75.2	2.5	467	13	US-10-027-632-297936	Sequence 297936, A
29	75.2	2.5	467	14	US-10-027-632-297935	Sequence 297935, A
30	75.2	2.5	467	14	US-10-027-632-297936	Sequence 297936, A
31	75.2	2.5	634	13	US-10-027-632-41737	Sequence 41737, A
32	75.2	2.5	634	13	US-10-027-632-41738	Sequence 41738, A
33	75.2	2.5	634	14	US-10-027-632-41737	Sequence 41737, A
34	75.2	2.5	634	14	US-10-027-632-41738	Sequence 41738, A
35	62	2.1	442	10	US-09-833-381-1965	Sequence 1965, Ap
36	59.2	2.0	469	13	US-10-027-632-62086	Sequence 62086, A
37	59.2	2.0	469	13	US-10-027-632-62087	Sequence 62087, A
38	59.2	2.0	469	13	US-10-027-632-63434	Sequence 63434, A
39	59.2	2.0	469	13	US-10-027-632-63435	Sequence 63435, A
40	59.2	2.0	469	14	US-10-027-632-62086	Sequence 62086, A
41	59.2	2.0	469	14	US-10-027-632-62087	Sequence 62087, A
42	59.2	2.0	469	14	US-10-027-632-63434	Sequence 63434, A
43	59.2	2.0	469	14	US-10-027-632-63435	Sequence 63435, A
44	57.6	1.9	2262	12	US-10-369-493-33072	Sequence 33072, A
45	56.4	1.9	344	10	US-09-960-352-4542	Sequence 4542, Ap

ALIGNMENTS

RESULT 1  
US-09-935-290-1  
; Sequence 1, Application US/09935290  
; Publication No. US20030044948A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller-Libermann, Rosana  
; TITLE OF INVENTION: 56919, A NOVEL HUMAN ACYLTRANSFERASE AND USES THEREOF  
; FILE REFERENCE: NMI-186  
; CURRENT APPLICATION NUMBER: US/09/935,290  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/226,509  
; PRIOR FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3003  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (341)...(2827)  
; NAME/KEY: misc feature  
; LOCATION: (1)...(3003)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-935-290-1

Query Match	99.9%	Score	3001	DB	11	Length	3003
Best Local Similarity	100.0%	Pred. No.	0	Mismatches	0	Indels	0
Matches	3003	Conservative	0				
Oy	1	TTCCGCCACCAAGGCTGCTGCGGGGGGACTCTTTCTGAGGTTACTGTGGAGCACCCAAAGTC	60				
Db	1	TTCCGCCACCAAGGCTGCTGCGGGGGGACTCTTTCTGAGGTTACTGTGGAGCACCCAAAGTC	60				
Oy	61	TGTCAGGCTCTGGCCGTGTCGCAACAGGACCCAGAGAACCCAGACCTGCTTATTCACCCA	120				
Db	61	TGTCAGGCTCTGGCCGTGTCGCAACAGGACCCAGAGAACCCAGACCTGCTTATTCACCCA	120				

Qy 121 CAGCTGGAGTGTCTTCCAGAGTCTCCATCAGCTTTTCTTAATCGACTGATGGAAAT 180  
Db 121 CAGCTGGAGTGTCTTCCAGAGTCTCCATCAGCTTTTCTTAATCGACTGATGGAAAT 180  
Qy 181 AATTCCTCAACACCAAGTCAAGATACAGGACGCGGTCCCTCTGTTGTATGGA 240  
Db 181 AATTCCTCAACACCAAGTCAAGATACAGGACGCGGTCCCTCTGTTGTATGGA 240  
Qy 241 CATTCCTGACCCGAACTGATAGTCTGAGTCTGAGTCTTATGTTATGAAACAGAGAAC 300  
Db 241 CATTCCTGACCCGAACTGATAGTCTGAGTCTTATGTTATGAAACAGAGAAC 300  
Qy 301 TTTTCATCCGACGACATGATTTGGGAATTACACTTTGTGACATGGATGAATCTGCACCTGAC 360  
Db 301 TTTTCATCCGACGACATGATTTGGGAATTACACTTTGTGACATGGATGAATCTGCACCTGAC 360  
Qy 361 CTTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCAGAAATACAGTCTTGGTGCATG 420  
Db 361 CTTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCAGAAATACAGTCTTGGTGCATG 420  
Qy 421 TAAGCACACAGTGAGGAATGGGTGAGTGTGGCTTTAGACCCACCGTCTTCAGATCTGC 480  
Db 421 TAAGCACACAGTGAGGAATGGGTGAGTGTGGCTTTAGACCCACCGTCTTCAGATCTGC 480  
Qy 481 AACTTTAAATGGAAGAAAGCCCTTAATGAGTCTGGAAGGCCATTTGTTGGAAGATGTTG 540  
Db 481 AACTTTAAATGGAAGAAAGCCCTTAATGAGTCTGGAAGGCCATTTGTTGGAAGATGTTG 540  
Qy 541 TTAATCTCTGACCTCCCGAGAGCTGGGACAAATTTTCAACCCAGTATCCCGTCTTTGGG 600  
Db 541 TTAATCTCTGACCTCCCGAGAGCTGGGACAAATTTTCAACCCAGTATCCCGTCTTTGGG 600  
Qy 601 TTTTGGCAATGTTATTTATATCAATGAATCTCACACAGACACCGCGGATGGCTTGAAG 660  
Db 601 TTTTGGCAATGTTATTTATATCAATGAATCTCACACAGACACCGCGGATGGCTTGAAG 660  
Qy 661 ACGCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTGCATAAGGGCATGTTTGGCAC 720  
Db 661 ACGCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTGCATAAGGGCATGTTTGGCAC 720  
Qy 721 CAATGTGACTGAAATGTGCTGAACACGACGTAGAGTACAAGAGCAATTCAGAAATGAC 780  
Db 721 CAATGTGACTGAAATGTGCTGAACACGACGTAGAGTACAAGAGCAATTCAGAAATGAC 780  
Qy 781 TGCTGAATTAACCTGATGGTCTGCCCGACGACCAATCAAGCCGTTAACAAGTGAA 840  
Db 781 TGCTGAATTAACCTGATGGTCTGCCCGACGACCAATCAAGCCGTTAACAAGTGAA 840  
Qy 841 AAAGAAAGCTAAAGGATCTTCAAGAAATGGTTGGCCACTGTCTCACCGGCAATGATCAG 900  
Db 841 AAAGAAAGCTAAAGGATCTTCAAGAAATGGTTGGCCACTGTCTCACCGGCAATGATCAG 900  
Qy 901 ACTGACTGGGTGGTGTCTGCTAAAATCTGTTCAACAGCTTCTTTTGGAAACATTCAAATTC 960  
Db 901 ACTGACTGGGTGGTGTCTGCTAAAATCTGTTCAACAGCTTCTTTTGGAAACATTCAAATTC 960  
Qy 961 CAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGACAGCAATTTGCGGCTCTGTTTCT 1020  
Db 961 CAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGACAGCAATTTGCGGCTCTGTTTCT 1020  
Qy 1021 ACCAGTTTCATAGATCCCATATTTGACTATCTGTCTCCTCATTTTCTCTCTCTGCCATAA 1080  
Db 1021 ACCAGTTTCATAGATCCCATATTTGACTATCTGTCTCCTCATTTTCTCTCTCTGCCATAA 1080  
Qy 1081 CATCAAGACCACTACATTTGCTTCAAGCAATAATCTCAACATCCCAATCTTCAGTACCTT 1140  
Db 1081 CATCAAGACCACTACATTTGCTTCAAGCAATAATCTCAACATCCCAATCTTCAGTACCTT 1140  
Qy 1141 GATCCATAAGCTTTGGGGCTTCTTCATACGACGAGCTCGATGAACACACAGATGGACG 1200  
Db 1141 GATCCATAAGCTTTGGGGCTTCTTCATACGACGAGCTCGATGAACACACAGATGGACG 1200  
Qy 1201 GAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATAGTTGAATTTACTTCGACAGCA 1260

1201 GAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATAGTTGAATTTACTTCGACAGCA 1260  
Qy 1261 GCATTTCTTGGAGATCTTCTCGAAGGCACAGTTCCTAGAGTGGAAACCTCTTGTGC 1320  
Db 1261 GCATTTCTTGGAGATCTTCTCGAAGGCACAGTTCCTAGAGTGGAAACCTCTTGTGC 1320  
Qy 1321 TCGGGCAGGACTTTTGTCTGAGTTGTGTAGATACTCTGTCTACCAATGTCTATCCAGACAT 1380  
Db 1321 TCGGGCAGGACTTTTGTCTGAGTTGTGTAGATACTCTGTCTACCAATGTCTATCCAGACAT 1380  
Qy 1381 CTTGTAATPACCTTGTGGAACTCTCTATGATCGCAATTCGAAAGTCTACTCAATGGTGA 1440  
Db 1381 CTTGTAATPACCTTGTGGAACTCTCTATGATCGCAATTCGAAAGTCTACTCAATGGTGA 1440  
Qy 1441 ACNACTGGGCAAACTTAAGAAGATGAGAGCTGTGGAGTGTAGCAGAGGTGTTATTAG 1500  
Db 1441 ACNACTGGGCAAACTTAAGAAGATGAGAGCTGTGGAGTGTAGCAGAGGTGTTATTAG 1500  
Qy 1501 AATGTTTACGAAATACTATGTTTGTCTCGAGTGGATTTTGCACAGCCATTTTCTTAA 1560  
Db 1501 AATGTTTACGAAATACTATGTTTGTCTCGAGTGGATTTTGCACAGCCATTTTCTTAA 1560  
Qy 1561 GGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCTCTACTTTCCCTGGAGCAAGC 1620  
Db 1561 GGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCTCTACTTTCCCTGGAGCAAGC 1620  
Qy 1621 GTTGTACAGCTATATCTTCTTCAAGACCCAGTGTGTCTGTCTACTTTCCCTGGAGCAC 1680  
Db 1621 GTTGTACAGCTATATCTTCTTCAAGACCCAGTGTGTCTGTCTACTTTCCCTGGAGCAC 1680  
Qy 1681 GTCCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGAGGTGATTCGAA 1740  
Db 1681 GTCCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGAGGTGATTCGAA 1740  
Qy 1741 TCTGGCTGAGCATATTTCTATTCACTGCTAGCAAGTCTCTGCTGCAATTTGTCACACAT 1800  
Db 1741 TCTGGCTGAGCATATTTCTATTCACTGCTAGCAAGTCTCTGCTGCAATTTGTCACACAT 1800  
Qy 1801 TGTGGCTGCTGCTCTCTACAGACAGGACAGGAATTTGATCTCTCACATTTGTGCA 1860  
Db 1801 TGTGGCTGCTGCTCTCTACAGACAGGACAGGAATTTGATCTCTCACATTTGTGCA 1860  
Qy 1861 AGACTTCTTTGTGATGAAAGAGAGTCTGCTGCTGATTTTGAACCTGGGTTCTCAGG 1920  
Db 1861 AGACTTCTTTGTGATGAAAGAGAGTCTGCTGCTGATTTTGAACCTGGGTTCTCAGG 1920  
Qy 1921 AAATTCAGAAAGATGTAGTAATGCATGCCATACAGCTGTGGGAAATTTGTGCACATC 1980  
Db 1921 AAATTCAGAAAGATGTAGTAATGCATGCCATACAGCTGTGGGAAATTTGTGCACATC 1980  
Qy 1981 CCACACTAGAGAAAGATGATTTTATACCCCCAGACAACTGTGCCATCAGTCTT 2040  
Db 1981 CCACACTAGAGAAAGATGATTTTATACCCCCAGACAACTGTGCCATCAGTCTT 2040  
Qy 2041 CGAACTCAACTCTACAGCAATGGGATCTCATGCTTTTATCATGGAGCCCATATAGC 2100  
Db 2041 CGAACTCAACTCTACAGCAATGGGATCTCATGCTTTTATCATGGAGCCCATATAGC 2100  
Qy 2101 TTTGACGCTTTTATGCAATTTCTGAAACAGAGGGAGTGGGGGTCCCATAGCACCCACC 2160  
Db 2101 TTTGACGCTTTTATGCAATTTCTGAAACAGAGGGAGTGGGGGTCCCATAGCACCCACC 2160  
Qy 2161 TAACTGTATCAGCAGGAGCAGCTGTGGGAGGGCGGCGCAGCTGTGTACCTTCTCTC 2220  
Db 2161 TAACTGTATCAGCAGGAGCAGCTGTGGGAGGGCGGCGCAGCTGTGTACCTTCTCTC 2220  
Qy 2221 CAATGAAGCACCATCTCACTGCTTGCAGACATTTTACCAGTCTGCCATGAACAGT 2280  
Db 2221 CAATGAAGCACCATCTCACTGCTTGCAGACATTTTACCAGTCTGCCATGAACAGT 2280  
Qy 2281 AGGAAAGTTTATCCAGTATGGCATTTCTTACAGTGGCAGACGATGACCAGGAATAT 2340



Db 783 CTGAATTAACCCCTGATGGTTCTGCCAGCAGCAATCAAAAGCGGTTAAACAAAGTGAAAA 842  
Qy 843 AGAAAGCTTAAAGGATTTCTCAAGAAATGGTTGCTCACTGCTCACCAGCAATGATCAGAC 902  
Db 843 AGAAAGCTTAAAGGATTTCTCAAGAAATGGTTGCTCACTGCTCACCAGCAATGATCAGAC 902  
Qy 903 TGACTGGGTGGGTGCTGCTTAAAACTGTTCAACAGCTTCTTTTGGAAATTTGCGGCTTCTGTTTCTAC 962  
Db 903 TGACTGGGTGGGTGCTGCTTAAAACTGTTCAACAGCTTCTTTTGGAAATTTCAAAATTCACA 962  
Qy 963 AAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACGAATTTGCGGCTTCTGTTTCTAC 1022  
Db 963 AAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACGAATTTGCGGCTTCTGTTTCTAC 1022  
Qy 1023 CAGTTTCATAGATCCCATATTTGACTATCTGCTGCTCACTTCACTTCTCTTTCGCCATAACA 1082  
Db 1023 CAGTTTCATAGATCCCATATTTGACTATCTGCTGCTCACTTCACTTCTCTTTCGCCATAACA 1082  
Qy 1083 TCAAGCACCATACATATGCTTTCAGGCAATATCTCAACATCCCAATCTTCAGTACCTTGA 1142  
Db 1083 TCAAGCACCATACATATGCTTTCAGGCAATATCTCAACATCCCAATCTTCAGTACCTTGA 1142  
Qy 1143 TCCATAAGCTTGGGGCTTCTTCATACGACGAAGGCTCGATGAAACACCAAGATGGACGGA 1202  
Db 1143 TCCATAAGCTTGGGGCTTCTTCATACGACGAAGGCTCGATGAAACACCAAGATGGACGGA 1202  
Qy 1203 AAGATGTTCTCTATAGAGCTTTGCTCCATGGCATATAGTTGAATTTACTTTCGACAGAGC 1262  
Db 1203 AAGATGTTCTCTATAGAGCTTTGCTCCATGGCATATAGTTGAATTTACTTTCGACAGAGC 1262  
Qy 1263 AATTCTTGGAGATCTTCTCGAAGGCACAGTTCTAGGAGTGAAAAACCTCTTTGTGCTC 1322  
Db 1263 AATTCTTGGAGATCTTCTCGAAGGCACAGTTCTAGGAGTGAAAAACCTCTTTGTGCTC 1322  
Qy 1323 GGGCAGGACTTTTGTCAAGTTGGTGTAGATATCTGCTACCAATGTCTATCCAGACATCT 1382  
Db 1323 GGGCAGGACTTTTGTCAAGTTGGTGTAGATATCTGCTACCAATGTCTATCCAGACATCT 1382  
Qy 1383 TGATATACCTGTGGATCTCTATGATCGCATATTCGAAGTCACTACAAATGGTGAAC 1442  
Db 1383 TGATATACCTGTGGATCTCTATGATCGCATATTCGAAGTCACTACAAATGGTGAAC 1442  
Qy 1443 AACTGGGCAAACTTAAGAAATGAGAGCTGTGGAGTGTAGCAAGAGGTTTATTAGAA 1502  
Db 1443 AACTGGGCAAACTTAAGAAATGAGAGCTGTGGAGTGTAGCAAGAGGTTTATTAGAA 1502  
Qy 1503 TGTTACGAAAAAATATGGTTGTGTCGAGTGGATTTTGACAGCCATTTTCTTAAAGG 1562  
Db 1503 TGTTACGAAAAAATATGGTTGTGTCGAGTGGATTTTGACAGCCATTTTCTTAAAGG 1562  
Qy 1563 AATATTTAGAAAGCAAGTCAGAACCGGTGTGCTCTACTTTCCCTGGAGCAAGCGT 1622  
Db 1563 AATATTTAGAAAGCAAGTCAGAACCGGTGTGCTCTACTTTCCCTGGAGCAAGCGT 1622  
Qy 1623 TGTTACAGCTATACCTTCAAGACCCAGTGTGCTGATGAAGGTAGAGACAGCT 1682  
Db 1623 TGTTACAGCTATACCTTCAAGACCCAGTGTGCTGATGAAGGTAGAGACAGCT 1682  
Qy 1683 CCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGGAGGTTGATTTGCAAAATC 1742  
Db 1683 CCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGGAGGTTGATTTGCAAAATC 1742  
Qy 1743 TGCTGAGCATATTTCTATTCATCTGCTAGCAAGTCTGTGGCATATGTCCACACACATTTG 1802  
Db 1743 TGCTGAGCATATTTCTATTCATCTGCTAGCAAGTCTGTGGCATATGTTCACACACATTTG 1802  
Qy 1803 TGCTTGCCTGCTCTCTACAGACACAGCAGGAAATTTGATCTCTCCACATTTGCTGAAG 1862  
Db 1803 TGCTTGCCTGCTCTCTACAGACACAGCAGGAAATTTGATCTCTCCACATTTGCTGAAG 1862  
Qy 1863 ACTTCTTTGTGATGAAGAGGAAGTCTTGGCTCGTGAATTTGACCTGGGGTCTTCAGGAA 1922  
Db 1863 ACTTCTTTGTGATGAAGAGGAAGTCTTGGCTCGTGAATTTGACCTGGGGTCTTCAGGAA 1922

Db 1863 ACTTCTTTGTGATGAAGAGGAAGTCTTGGCTCGTGAATTTGACCTGGGTTCTCAGGAA 1922  
Qy 1923 ATTCAGAAAGATGTAGTAATGTCATGTCATATGAGCTGTGCGAAATTTGTGTACAAATCACCC 1982  
Db 1923 ATTCAGAAAGATGTAGTAATGTCATGTCATATGAGCTGTGCGAAATTTGTGTACAAATCACCC 1982  
Qy 1983 ACATAAGCAGGAACGATGATGTTTATCACCCTCCAGCAACATGCTGCCATCAGTCTTCG 2042  
Db 1983 ACATAAGCAGGAACGATGATGTTTATCACCCTCCAGCAACATGCTGCCATCAGTCTTCG 2042  
Qy 2043 AACTCAACTTCTACAGCAATGGGGTACTTCAATGCTTTTATCATGAGGAGCATCATAGCTT 2102  
Db 2043 AACTCAACTTCTACAGCAATGGGGTACTTCAATGCTTTTATCATGAGGAGCATCATAGCTT 2102  
Qy 2103 GCAGCCTTTATGAGGAGGAGTCTGCGGAAAGCGCCAGCCTGTGCTTCTCTCCA 2222  
Db 2103 GCAGCCTTTATGAGGAGGAGTCTGCGGAAAGCGCCAGCCTGTGCTTCTCTCCA 2222  
Qy 2223 ATGAAGGCAACCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAAACAGTAG 2282  
Db 2223 ATGAAGGCAACCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAAACAGTAG 2282  
Qy 2283 GAAAGTTTATCAGTATGAGTGGGCAAGAGCTTCCAGAACCTTTGTCTTTGGAGAA 2342  
Db 2283 GAAAGTTTATCAGTATGAGTGGGCAAGAGCTTCCAGAACCTTTGTCTTTGGAGAA 2342  
Qy 2343 GTCCTAGTCTTGTGAGCAGTGGGCAAGAGCTTCCAGAACCTTTGTCTTTGGAGAA 2402  
Db 2343 GTCCTAGTCTTGTGAGCAGTGGGCAAGAGCTTCCAGAACCTTTGTCTTTGGAGAA 2402  
Qy 2403 GTGATGAAGAAATGAGTGAAGTGTGTTGGGAGGAAACAGGAGATGCTACTGAAGG 2462  
Db 2403 GTGATGAAGAAATGAGTGAAGTGTGTTGGGAGGAAACAGGAGATGCTACTGAAGG 2462  
Qy 2463 TGAGCCAAATCCAGGAGCACCAGAGTGTATCACCCTTTTACAGAGACTCTTTGGGCTT 2522  
Db 2463 TGAGCCAAATCCAGGAGCACCAGAGTGTATCACCCTTTTACAGAGACTCTTTGGGCTT 2522  
Qy 2523 TGCTGAGGCTTACAGCTCTGCTGCTTGTGTTTACAACTTTCAGTGTGCTGTTCCAG 2582  
Db 2523 TGCTGAGGCTTACAGCTCTGCTGCTTGTGTTTACAACTTTCAGTGTGCTGTTCCAG 2582  
Qy 2583 AACCTGAGTATCTGCAAAAAGTTGCAAAAATCCTTAATAACCAAGAAAGAAATGTTG 2642  
Db 2583 AACCTGAGTATCTGCAAAAAGTTGCAAAAATCCTTAATAACCAAGAAAGAAATGTTG 2642  
Qy 2643 CAGTATATGCTGAGAGTGCACATATTTGTTGTGAAGAAATGCTGTGAAATGTTTAAAGG 2702  
Db 2643 CAGTATATGCTGAGAGTGCACATATTTGTTGTGAAGAAATGCTGTGAAATGTTTAAAGG 2702  
Qy 2703 ATATTGGGGTTTTCAAGGAGACCAACAAAGAGAGTGTCTGTTTTAGAACTTGAGCAGCA 2762  
Db 2703 ATATTGGGGTTTTCAAGGAGACCAACAAAGAGAGTGTCTGTTTTAGAACTTGAGCAGCA 2762  
Qy 2763 CTTTCTTACCTCAATGCAACCGAACAAAATCTTCTAGAATATATTTCTGAGTTTGTGGTGC 2822  
Db 2763 CTTTCTTACCTCAATGCAACCGAACAAAATCTTCTAGAATATATTTCTGAGTTTGTGGTGC 2822  
Qy 2823 TGTAGTAACTGTGGCAGTGTGCGCAATGAAGTCAAGATGAGTGTGCTTTGTTAGGTA 2882  
Db 2823 TGTAGTAACTGTGGCAGTGTGCGCAATGAAGTCAAGTGTGCTTTGTTAGGTA 2882  
Qy 2883 CCAGCTTCTGGCTCAAGAGTGTGCGCAAGGAGTGTGCGGAGTGTGCTTTGTTAGGTA 2942  
Db 2883 CCAGCTTCTGGCTCAAGAGTGTGCGCAAGGAGTGTGCGGAGTGTGCTTTGTTAGGTA 2942  
Qy 2943 AAGTGAATCTCTGGAAGCAAGTGTCTTCTNCTCCATGGATCTGAGATCTTTCCAGCTT 3002  
Db 2943 AAGTGAATCTCTGGAAGCAAGTGTCTTCTNCTCCATGGATCTGAGATCTTTCCAGCTT 3002

QY 3003 T 3003  
Db 3001 T 3001

## RESULT 3

US-10-288-252-26

; Sequence 26, Application US/10288252

; Publication No. US20030143686A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: LAL, Preeti G.

; APPLICANT: TANG, Y. Tom

; APPLICANT: YUE, Henry

; APPLICANT: BURFORD, Neil

; APPLICANT: GANDHI, Ameena R.

; APPLICANT: WARREN, Bridget A.

; APPLICANT: YAO, Monique G.

; APPLICANT: TRIBOULEY, Catherine M.

; APPLICANT: BAUGHN, Mariah R.

; APPLICANT: LEE, Ernestine A.

; APPLICANT: HAFALIA, April J.A.

; APPLICANT: LU, Yan

; APPLICANT: GRIFFIN, Jennifer A.

; APPLICANT: SANJANWALA, Madhu S.

; APPLICANT: DING, Li

; TITLE OF INVENTION: TRANSFERASES

; FILE REFERENCE: PI-0241 USA

; CURRENT APPLICATION NUMBER: US/10/288,252

; CURRENT FILING DATE: 2002-11-04

; PRIOR APPLICATION NUMBER: PCT US01/30424

; PRIOR FILING DATE: 2001-09-28

; PRIOR APPLICATION NUMBER: US 60/252,819

; PRIOR FILING DATE: 2000-11-21

; PRIOR APPLICATION NUMBER: US 60/249,639

; PRIOR FILING DATE: 2000-11-16

; PRIOR APPLICATION NUMBER: US 60/247,931

; PRIOR FILING DATE: 2000-11-09

; PRIOR APPLICATION NUMBER: US 60/246,001

; PRIOR FILING DATE: 2000-11-03

; PRIOR APPLICATION NUMBER: US 60/244,025

; PRIOR FILING DATE: 2000-10-27

; PRIOR APPLICATION NUMBER: US 60/238,481

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: US 60/236,523

; PRIOR FILING DATE: 2000-09-29

; NUMBER OF SEQ ID NOS: 40

; SOFTWARE: PERL Program

; SEQ ID NO 26

; LENGTH: 3210

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20030143686A1 1517230CB1

US-10-288-252-26

Query Match 98.5%; Score 2958; DB 13; Length 3210;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2983; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 12 GCTGCTGGGGGGGACTCTTTCTGAGGTTACTGTGGAGCACCACCAAGTCTGTGAGCCCTCT 71  
Db 9 GCTGCTGGGGGGGACTCTTTCTGAGGTTACTGTGGAGCACCACCAAGTCTGTGAGCCCTCT 68

QY 72 GGCCTGTCAAAACAGGACCCAGGAGAACCCAGACCTTGTCTTAATTCACCCACAGCCCTGGGAC 131  
Db 69 GGCCTGTCAAAACAGGACCCAGGAGAACCCAGACCTTGTCTTAATTCACCCACAGCCCTGGGAC 128

QY 132 TGTCTTCTCCAGAGTCTCATCAGCTTGTGCTTAATTCGACTGATTTGAAATAATTTCTTCAA 191  
Db 129 TGTCTTCTCCAGAGTCTCATCAGCTTGTGCTTAATTCGACTGATTTGAAATAATTTCTTCAA 188

QY 1272 AGATCTTCTGGAGGACACAGTCTTAGAGTGGAAAAACCTCTGTGCTCGGCGAGGAC 1331  
Db |||||  
QY 1269 AGATCTTCTGGAGGACACAGTCTTAGAGTGGAAAAACCTCTGTGCTCGGCGAGGAC 1328  
Db |||||  
QY 1332 TTTTGTCAAGTGGTAGACTCTGTCTACCAATGTCTATCCAGACATCTTGTATATAC 1391  
Db |||||  
QY 1329 TTTTGTCAAGTGGTAGACTCTGTCTACCAATGTCTATCCAGACATCTTGTATATAC 1388  
Db |||||  
QY 1392 CTGTGTGAATCTCTATGATCGCATATCGAAGTCACTACAATGGTGAACAACCTGGGCA 1451  
Db |||||  
QY 1389 CTGTGTGAATCTCTATGATCGCATATCGAAGTCACTACAATGGTGAACAACCTGGGCA 1448  
Db |||||  
QY 1452 AACCTAAGAAAGATGAGAGCCTGTGGAGTGTAGCAAGAGGTGTATTTAGAAATGTACGAA 1511  
Db |||||  
QY 1449 AACCTAAGAAAGATGAGAGCCTGTGGAGTGTAGCAAGAGGTGTATTTAGAAATGTACGAA 1508  
Db |||||  
QY 1512 AAAAATATGTTGTGTCGAGTGGATTTTGCACAGCCATTTTCTTTAAAGGAATATTAG 1571  
Db |||||  
QY 1509 AAAAATATGTTGTGTCGAGTGGATTTTGCACAGCCATTTTCTTTAAAGGAATATTAG 1568  
Db |||||  
QY 1572 AAAGCCAAAGTCAGAAACCGGTGTCTCTACTTTTCCCTGGAGCAAGCGTGTGTACCAG 1631  
Db |||||  
QY 1569 AAAGCCAAAGTCAGAAACCGGTGTCTCTACTTTTCCCTGGAGCAAGCGTGTGTACCAG 1628  
Db |||||  
QY 1632 CTATACTTCTTCAAGACCCAGTGTCTCTGATGAAGGTAGAGACACGTCCATTAATG 1691  
Db |||||  
QY 1629 CTATACTTCTTCAAGACCCAGTGTCTCTGATGAAGGTAGAGACACGTCCATTAATG 1688  
Db |||||  
QY 1692 AGTCCAGAAATGCAACAGATGAATCCCTAGCAAGGAGGTGTATGCAATCTGGCTGAGC 1751  
Db |||||  
QY 1689 AGTCCAGAAATGCAACAGATGAATCCCTAGCAAGGAGGTGTATGCAAAATCTGGCTGAGC 1748  
Db |||||  
QY 1752 ATATTCTATTCTAGTGTAGCAAGTCTGTGCGCAATTTGTCACACACATTTGGCTGTGCC 1811  
Db |||||  
QY 1749 ATATTCTATTCTAGTGTAGCAAGTCTGTGCGCAATTTGTCACACACATTTGGCTGTGCC 1808  
Db |||||  
QY 1812 TGCTCTCTCTACAGACACAGCAGGGAATGATCTCTCCATTTGGTGTGCAAGACTTCTTTG 1871  
Db |||||  
QY 1809 TGCTCTCTCTACAGACACAGCAGGGAATGATCTCTCCATTTGGTGTGCAAGACTTCTTTG 1868  
Db |||||  
QY 1872 TGATGAAGGAAGTCTGGCTGTGATTTTGAACCTGGGGTCTCAGGAAATTCAGAAG 1931  
Db |||||  
QY 1869 TGATGAAGGAAGTCTGGCTGTGATTTTGAACCTGGGGTCTCAGGAAATTCAGAAG 1928  
Db |||||  
QY 1932 ATCTAGTAAATGCATGCATACAGTCTGTGGAAATTTGTGTCACAAATCACCACACTAGCA 1991  
Db |||||  
QY 1929 ATCTAGTAAATGCATGCATACAGTCTGTGGAAATTTGTGTCACAAATCACCACACTAGCA 1988  
Db |||||  
QY 1992 GGAACGATGAGTTTTTTATCACCCCGACAGCAACTGTCCCATCAGTCTTCGAACTCAACT 2051  
Db |||||  
QY 1989 GGAACGATGAGTTTTTTATCACCCCGACAGCAACTGTCCCATCAGTCTTCGAACTCAACT 2048  
Db |||||  
QY 2052 TCTACAGCAATGGGGTACTTCTATGCTTTTATCATATGAGGCGCATCATAGCTTGCAGCCTTT 2111  
Db |||||  
QY 2049 TCTACAGCAATGGGGTACTTCTATGCTTTTATCATATGAGGCGCATCATAGCTTGCAGCCTTT 2108  
Db |||||  
QY 2112 ATTCAGTCTTGAACAGAGGGGACTGGGGGTCCCACTAGCACCCACCTAACCTCATCA 2171  
Db |||||  
QY 2109 ATTCAGTCTTGAACAGAGGGGACTGGGGGTCCCACTAGCACCCACCTAACCTCATCA 2168  
Db |||||  
QY 2172 GCCAGGAGCAGTGTGGTGGGAGCGGCGCAGCGCTGTGCTACTTCTTCCAATGAAGGCA 2231  
Db |||||  
QY 2169 GCCAGGAGCAGTGTGGTGGGAGCGGCGCAGCGCTGTGCTACTTCTTCCAATGAAGGCA 2228  
Db |||||  
QY 2232 CCATCTCACTGCTTGGCAGACATTTTATCAAGTGTGCCATGAAACAGTAGGAAAGTTTA 2291  
Db |||||  
QY 2229 CCATCTCACTGCTTGGCAGACATTTTATCAAGTGTGCCATGAAACAGTAGGAAAGTTTA 2288  
Db |||||  
QY 2292 TCCAGTATGCAATTTCTACAGTGGCAGAGCACCATGACCCAGGAGATATCAGTCTTAGTC 2351  
Db |||||  
QY 2289 TCCAGTATGCAATTTCTACAGTGGCAGAGCACCATGACCCAGGAGATATCAGTCTTAGTC 2348  
Db |||||  
QY 2352 TTGCTGAGCAGCAGTGGGCAAGAAGCTTCCAGAACTTTGTTCTTGGAGAAAGTGATGAAG 2411  
Db |||||

Db 2349 TTGCTGAGCAGCAGTGGGCAAGAAGCTTCTTGAACCTTTGTCTTGGAGAAGTGATGAAG 2408  
QY |||||  
Db 2412 AAGATCAAGACAGTGTACTTTGGGGAGGAACAGCAGATTTGCTACCTGAAAGGTGAGCAAT 2471  
QY |||||  
Db 2409 AAGATCAAGACAGTGTACTTTGGGGAGGAACAGCAGATTTGCTACCTGAAAGGTGAGCAAT 2468  
QY |||||  
Db 2472 CCAAGGAGCACCAGCAGTGTATACACCTTCTTACAGAGACTCTCTTGGGCCCTTTGCTGAGG 2531  
QY |||||  
Db 2469 CCAAGGAGCACCAGCAGTGTATACACCTTCTTACAGAGACTCTCTTGGGCCCTTTGCTGAGG 2528  
QY |||||  
Db 2532 CTTACAGCTCTGTGTCATCTTTGTTTCAAACTTCAAGTGGTCTGTTCAGAAACCTGAGT 2591  
QY |||||  
Db 2529 CTTACAGCTCTGTGTCATCTTTGTTTCAAACTTCAAGTGGTCTGTTCAGAAACCTGAGT 2588  
QY |||||  
Db 2592 ATCTGCAAAAGTTGCAAAATACCTTAATAACAGAAACAGAAAGAAATGTTGCAAGTATATG 2651  
QY |||||  
Db 2589 ATCTGCAAAAGTTGCAAAATACCTTAATAACAGAAACAGAAAGAAATGTTGCAAGTATATG 2648  
QY |||||  
Db 2652 CTGAGAGTCCCATATATTGTTGTAAGAAATGCTGTGAAATGTTTAAAGGATATTGGGG 2711  
QY |||||  
Db 2649 CTGAGAGTCCCATATATTGTTGTAAGAAATGCTGTGAAATGTTTAAAGGATATTGGGG 2708  
QY |||||  
Db 2712 TTTTCAAGGAGACCAAAACAAAGAGAGTGTCTGTTTTAGAACTGAGCAGCACTTTTCTAC 2771  
QY |||||  
Db 2709 TTTTCAAGGAGACCAAAACAAAGAGAGTGTCTGTTTTAGAACTGAGCAGCACTTTTCTAC 2768  
QY |||||  
Db 2772 CTCATATGCAACCGACAAACAACTTCTAGAAATATATCTGAGTGTGTTGCTGTAGGTAA 2831  
QY |||||  
Db 2769 CTCATATGCAACCGACAAACAACTTCTAGAAATATATCTGAGTGTGTTGCTGTAGGTAA 2828  
QY |||||  
Db 2832 CGTGTGGCACTGTGCGCAATGAAGTCTGATGAGTGTCTTGTAGGTACCAGCTTCT 2891  
QY |||||  
Db 2829 CGTGTGGCACTGTGCGCAATGAAGTCTGATGAGTGTCTTGTAGGTACCAGCTTCT 2888  
QY |||||  
Db 2892 GGTCTAAGAGTTGAAGTGTCTTCCAGGGGTGAGGCTGCGCTGTGCTGCGAAAGTGTCT 2951  
QY |||||  
Db 2889 GGTCTAAGAG-ITGAAGGTGCGTGTGCA-GGGTGAGGCTGCGCTGTGCTGCGAAAGTGTCT 2946  
QY |||||  
Db 2952 CTTGGAAGCAAGTGTCTTCCCTCCATGGAATCTGAGATCTTCCAGCTTT 3003  
QY |||||  
Db 2947 CTTGGAAGCAAGTGTCTTCCCTCCATGGAATCTGATCTTCCAGCTTCT 2998  
QY |||||

## RESULT 4

US-09-935-290-3  
; Sequence 3, Application US/09935290  
; Publication No. US2003004948A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller-Libermann, Rosana  
; TITLE OF INVENTION: 56919, A NOVEL HUMAN ACYLTRANSFERASE AND USES THEREOF  
; FILE REFERENCE: MNI-186  
; CURRENT APPLICATION NUMBER: US/09/935,290  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/226,509  
; PRIOR FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2484  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-935-290-3

Query Match 82.7%; Score 2484; DB 11; Length 2484;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 341 ATGGATGAATCTGCACTGACCTTGGTCAATAGATGTTTCTTATCTGCCACATTCATCA 400  
Db 1 ATGGATGAATCTGCACTGACCTTGGTCAATAGATGTTTCTTATCTGCCACATTCATCA 60  
QY 401 GAATACAGTGTGTTGTCGATGTAAGCACACAGTAGGAAATGGGGTGAAGTGTGGCTTTAGA 460



Db 61 GAATACAGTGGTGGTGAATGAAGCACACAGTGAAGGAATGGGGTGAAGTGGCTTTTGA 120  
Qy  
Db 461 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAATGAGTCGGAAGG 520  
Qy  
Db 121 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAATGAGTCGGAAGG 180  
Qy  
Db 521 CCAATTTGTTGAAGATGTTGTTATCTCTGCACTCCGCCAGAGCTGGGACAAATTTTCAAC 580  
Qy  
Db 181 CCAATTTGTTGAAGATGTTGTTATCTCTGCACTCCGCCAGAGCTGGGACAAATTTTCAAC 240  
Qy  
Db 581 CCCAGTATCCCGTCTTTGGGTTGGGAATGTTATTTATATCAATGAAGTCAACACAGA 640  
Qy  
Db 241 CCCAGTATCCCGTCTTTGGGTTGGGAATGTTATTTATATCAATGAAGTCAACACAGA 300  
Qy  
Db 641 CACCGGATGGCTTTCGAAGCGCTTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTG 700  
Qy  
Db 301 CACCGGATGGCTTTCGAAGCGCTTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTG 360  
Qy  
Db 701 CATAGGCGATGTTTGGCCACCAATGTGACTGAAATGTGCTGAAAGAGTGAAGTGAAGTCA 760  
Qy  
Db 361 CATAGGCGATGTTTGGCCACCAATGTGACTGAAATGTGCTGAAAGAGTGAAGTGAAGTCA 420  
Qy  
Db 761 GAGCAATTCAGAGTGGCTGCTGAATTAACCTGATGGTCTGCCCCAGCAGCAATCA 820  
Qy  
Db 421 GAGCAATTCAGAGTGGCTGCTGAATTAACCTGATGGTCTGCCCCAGCAGCAATCA 480  
Qy  
Db 821 AAAGCCGTTAAAGAGTGAAGAAAGAAAGCTTAAAGGAATCTTCAAGAAATGTTGCCACT 880  
Qy  
Db 481 AAAGCCGTTAAAGAGTGAAGAAAGAAAGCTTAAAGGAATCTTCAAGAAATGTTGCCACT 540  
Qy  
Db 881 GTCTCACCGCAATGATCAGACTGCTGGGTGGTGTCTGCTAAACCTGTTCAACAGCTTC 940  
Qy  
Db 541 GTCTCACCGCAATGATCAGACTGCTGGGTGGTGTCTGCTAAACCTGTTCAACAGCTTC 600  
Qy  
Db 941 TTTTGGAACTTCAAAATTCAGAGTCAAAAGTCAAAAGTCAAAAGTCAAAAGTCAAAAGTCA 1000  
Qy  
Db 601 TTTTGGAACTTCAAAATTCAGAGTCAAAAGTCAAAAGTCAAAAGTCAAAAGTCAAAAGTCA 660  
Qy  
Db 1001 AATTTGGCGTCTCTGTTTCTACAGTTCATAGATCCCATATTTGATATCTGCTGCTCACT 1060  
Qy  
Db 661 AATTTGGCGTCTCTGTTTCTACAGTTCATAGATCCCATATTTGATATCTGCTGCTCACT 720  
Qy  
Db 1061 TTCAATCTCTTCTGCAATTAACATCAAGCACCACATATTTGCTTCAAGCAATTAATCTCAAC 1120  
Qy  
Db 721 TTCAATCTCTTCTGCAATTAACATCAAGCACCACATATTTGCTTCAAGCAATTAATCTCAAC 780  
Qy  
Db 1121 ATCCCAATCTTCAAGTACCTTGAATCAATGAAGCTTGGGGCTTCTTCAACGAGAGGCTC 1180  
Qy  
Db 781 ATCCCAATCTTCAAGTACCTTGAATCAATGAAGCTTGGGGCTTCTTCAACGAGAGGCTC 840  
Qy  
Db 1181 GATGAACACACAGATGGAGCAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1240  
Qy  
Db 841 GATGAACACACAGATGGAGCAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Qy  
Db 1241 GTTGAATTTACTTCGACAGAGCAATTTCTTGGAGATCTTCTGGAAGGCAACGTTCTAGG 1300  
Qy  
Db 901 GTTGAATTTACTTCGACAGAGCAATTTCTTGGAGATCTTCTGGAAGGCAACGTTCTAGG 960  
Qy  
Db 1301 AGTGAAGAAACCTCTTGTGCTCGGCAGGAATTTTGTGAGTGTGGTATGATCTGTGCT 1360  
Qy  
Db 961 AGTGAAGAAACCTCTTGTGCTCGGCAGGAATTTTGTGAGTGTGGTATGATCTGTGCT 1020  
Qy  
Db 1361 ACCAATGTCATCCAGACATCTGATATATACTGTTGAATCTCTATGATCGCATATC 1420  
Qy  
Db 1021 ACCAATGTCATCCAGACATCTGATATATACTGTTGAATCTCTATGATCGCATATC 1080  
Qy  
Db 1421 GAAGGTCACTAATGTTGAACAACTGGGCAAACTTAAGAAATGAAGCTGTGGAGT 1480  
Qy  
Db 1081 GAAGGTCACTAATGTTGAACAACTGGGCAAACTTAAGAAATGAAGCTGTGGAGT 1140  
Qy  
Db 1481 GTAGCAAGAGTGTATTAGAAATGTTACGAAAAAATCTATGTTGTGTCGAGTGGATTTT 1540

Db 1141 GTAGCAAGAGTGTATTAGAAATGTTACGAAAAAATCTATGGTGTGTCGAGTGGATTTT 1200  
Qy  
Db 1541 GCACAGCCATTTTCTCTTAAAGGAATATTAGAAAGCAAAAGTCAAGAACCGGTGCTGCT 1600  
Qy  
Db 1201 GCACAGCCATTTTCTCTTAAAGGAATATTAGAAAGCAAAAGTCAAGAACCGGTGCTGCT 1260  
Qy  
Db 1601 CTACTTTCCCTGGAGCAAGCGTTGTTACAGCTATATCTTCTTCAAGACCCAGTGAATGCT 1660  
Qy  
Db 1261 CTACTTTCCCTGGAGCAAGCGTTGTTACAGCTATATCTTCTTCAAGACCCAGTGAATGCT 1320  
Qy  
Db 1661 GCTGATCAAGGTGAGACAGCGTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1720  
Qy  
Db 1321 GCTGATCAAGGTGAGACAGCGTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Qy  
Db 1721 CGAAGAGGTTGATTGTCAAATCTGCTGAGCATATTTATTTACTGTCTAGCAAGTCCCTGT 1780  
Qy  
Db 1381 CGAAGAGGTTGATTGTCAAATCTGCTGAGCATATTTATTTACTGTCTAGCAAGTCCCTGT 1440  
Qy  
Db 1781 GCCAATTATGTCACACATTTGTGGCTTGTCTCTCTCTACAGACACAGCGGGAAT 1840  
Qy  
Db 1441 GCCAATTATGTCACACATTTGTGGCTTGTCTCTCTCTACAGACACAGCGGGAAT 1500  
Qy  
Db 1841 GATCTCTCCACATTTGTCGAAGACTTCTTTGTGATGAAGAGGAAAGTCTTGGCTCGTGA 1900  
Qy  
Db 1501 GATCTCTCCACATTTGTCGAAGACTTCTTTGTGATGAAGAGGAAAGTCTTGGCTCGTGA 1560  
Qy  
Db 1901 TTTGACCTGGGTTCTCAGGAAATTCAGAAAGTGTAGTAAATGATGTCATGTCATACAGTCTG 1960  
Qy  
Db 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAAGTGTAGTAAATGATGTCATGTCATACAGTCTG 1620  
Qy  
Db 1961 GGAAATTTGTGTCAAAATCAACCCACATAGCAGGAAAGTGTATTTTATCACCCCCAGC 2020  
Qy  
Db 1621 GGAAATTTGTGTCAAAATCAACCCACATAGCAGGAAAGTGTATTTTATCACCCCCAGC 1680  
Qy  
Db 2021 ACAACTGTCCCATCAGTCTTCCGAACTCAACTTCTACAGCAATGGGTAATCTTCTGTTT 2080  
Qy  
Db 1681 ACAACTGTCCCATCAGTCTTCCGAACTCAACTTCTACAGCAATGGGTAATCTTCTGTTT 1740  
Qy  
Db 2081 ATCAATGAGGCAATCATAGCTTTCAGAGTTCATGAGTTCGAAACAGAGGGGACTGGG 2140  
Qy  
Db 1741 ATCAATGAGGCAATCATAGCTTTCAGAGTTCATGAGTTCGAAACAGAGGGGACTGGG 1800  
Qy  
Db 2141 GGTCCCACTAGCACCCCACTTAACCTGATCAGCAGGAGCAGCTGTGCGGAAGCGCGCC 2200  
Qy  
Db 1801 GGTCCCACTAGCACCCCACTTAACCTGATCAGCAGGAGCAGCTGTGCGGAAGCGCGCC 1860  
Qy  
Db 2201 AGCTGTGCTACCTTCTTCCATGAAGGACCACTCTCACTGCTTGCAGACATTTTAC 2260  
Qy  
Db 1861 AGCTGTGCTACCTTCTTCCATGAAGGACCACTCTCACTGCTTGCAGACATTTTAC 1920  
Qy  
Db 2261 CAAGTCTGCCATGAACACAGTAGGAAGTTCATCCAGTATGGCATTTCTTACAGTGGCAGAG 2320  
Qy  
Db 1921 CAAGTCTGCCATGAACACAGTAGGAAGTTCATCCAGTATGGCATTTCTTACAGTGGCAGAG 1980  
Qy  
Db 2321 CACGATGACCAAGGAAGATATCAGTCTTCTGCTGAGCAGCAGTGGGCAAGAAAGCTT 2380  
Qy  
Db 1981 CACGATGACCAAGGAAGATATCAGTCTTCTGCTGAGCAGCAGTGGGCAAGAAAGCTT 2040  
Qy  
Db 2381 CCAGAACTTTTGTCTTGGAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAG 2440  
Qy  
Db 2041 CCAGAACTTTTGTCTTGGAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAG 2100  
Qy  
Db 2441 CAGCGAGATGCTTACCTGAAGGTGAGCCAAATCCAAAGGAGCAGCAGAGTTCACCTTC 2500  
Qy  
Db 2101 CAGCGAGATGCTTACCTGAAGGTGAGCCAAATCCAAAGGAGCAGCAGAGTTCACCTTC 2160  
Qy  
Db 2501 TTACAGAGATCTCTTGGGCTTGTCTGAGGCTTACAGTCTCTGCTGCTCTTGTGTTTAC 2560  
Qy  
Db 2161 TTACAGAGATCTCTTGGGCTTGTCTGAGGCTTACAGTCTCTGCTGCTCTTGTGTTTAC 2220  
Qy  
Db 2561 AACTTCAAGTGGTCTGTTCCAGAACTGAGTATCTGCAAAAGTGTGCAAAATACCTAATA 2620  
Qy  
Db 2221 AACTTCAAGTGGTCTGTTCCAGAACTGAGTATCTGCAAAAGTGTGCAAAATACCTAATA 2280

QY 2621 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATGCTTTGTGAAG 2680  
DB |||||  
QY 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATGCTTTGTGAAG 2340  
DB |||||  
QY 2681 AATGCTGTGAAATGTTTAAGATATTTGGGGTTTCAAGGAGACCAAAACAAAGAGAGTG 2740  
DB |||||  
QY 2341 AATGCTGTGAAATGTTTAAGATATTTGGGGTTTCAAGGAGACCAAAACAAAGAGAGTG 2400  
DB |||||  
QY 2741 TCTGTTTTAGAACTGAGCAGCACCTTTTCTACCTCAATGCAACCGACAAACAACTTCTAGAA 2800  
DB |||||  
QY 2401 TCTGTTTTAGAACTGAGCAGCACCTTTTCTACCTCAATGCAACCGACAAACAACTTCTAGAA 2460  
DB |||||  
QY 2801 TATATTTCTGAGTTTGTGTGCTG 2824  
DB |||||  
QY 2461 TATATTTCTGAGTTTGTGTGCTG 2484  
DB |||||

## RESULT 5

US-09-917-800A-1648  
; Sequence 1648, Application US/09917800A  
; Patent No. US20020119462A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendrick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Castle, Arthur  
; APPLICANT: Elashoff, Michael  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCE: 44921-5038-US  
; CURRENT APPLICATION NUMBER: US/09/917,800A  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,040  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,880  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: US 60/290,029  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US 60/290,645  
; PRIOR FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: US 60/292,336  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/295,798  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/297,457  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,884  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,459  
; PRIOR FILING DATE: 2001-07-09  
; NUMBER OF SEQ ID NOS: 1740  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1648  
; LENGTH: 2646  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM\_017274  
US-09-917-800A-1648

Query Match 63.1%; Score 1896.2; DB 10; Length 2646;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2126; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

QY 319 TTTGGGAATTACACTTTGTGACATGGATGAATCTGCACCTGACCTTTGGTACAATAGATGT 378  
DB |||||  
QY 138 TTTGGGACTTCCACGTTGAGTCATGGAGAGTCTTCAGTGACAATTTGGCAATAGACGT 197  
DB |||||  
QY 379 TTCTTATCTGCACATTCATCAGATACAGTGTGGTCGATGTAAGCACAAGAGTGAGGA 438  
DB |||||  
QY 198 TTCTTATCTGCCCAATTCATCAGAAATACAGCCCTTGGCCGATGTAACACACGAATGAGGA 257  
DB |||||

QY 439 ATGGGGTGAGTGTGGCTTTAGACCCACCGCTCTTTCAGATCTGCAACTTTAAATGGAAGA 498  
DB |||||  
QY 258 CTGGGTTGACTGTGGCTTCAAAACCTTACCTTCTTCAGATCCGCAACCGTGAATGGAAGGA 317  
DB |||||  
QY 499 AAGCCTAATGAGTCGAAAAGGCCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCCA 558  
DB |||||  
QY 318 GAGCCTCATGAGCCGGAAGAGGCCCTTCTGTTGGAAGGTTGCTATTTCATGCAACGCTCA 377  
DB |||||  
QY 559 GAGCTGGGACAAATTTTCAACCCAGTATCCGCTCTTTGGGTTTTCGGAATGTTTATTTA 618  
DB |||||  
QY 378 GAGCTGGGAAGAGTTTTCACCCAGTATCCATCTCTGGGTTTTCGGAATGTTTATTTA 437  
DB |||||  
QY 619 TATCAATGAACTCACACAAGACACCGCGGATGGCTTGAAGAAGCGCTTTCTTACGTTCT 678  
DB |||||  
QY 438 TATCAATGAACTCACACAAGGACCGGAGTGGCTGGCAAGAGAGCGCTTTCTTACATCCT 497  
DB |||||  
QY 679 TTTTATTCAAGAGCGAGATGTGCATGAAGGCGATGTTTGGCCACCAATGTGACTGAAATGT 738  
DB |||||  
QY 498 TTTTGTTCAGAGCGCGATGTCCACAAGGGCATGTTTGGCCACCGATTCACGTACAATGT 557  
DB |||||  
QY 739 GCTGAACAGCAGTAGAGTACAAGAGCAATTCAGAAAGTGGCTGTGTAATTTAAACCTGA 798  
DB |||||  
QY 558 ACTGAATAGCAGAGATCCAAAGAGCAATTCCTGAAGTGGCTGCAGAAATGAAACCGGA 617  
DB |||||  
QY 799 TGGTTCTGCCAGCAGCAATCAAAAGCCGTTTAACAAAGTGAAGAAAGAGCTTAAAGAGAT 858  
DB |||||  
QY 618 TGGATCTGCCAGCAGCAGTCCAAAGCCATCCAGAAAGTGAAGAGCAAGGAGGAGAT 677  
DB |||||  
QY 859 TCTTCAAGAAATGTTGGCCACTGTCTCACCGGCAATGATCAGACTGACTGGGTGGTGTCT 918  
DB |||||  
QY 678 CTTCCAGGAAATGTTGTGTACAGTCTTCCCGGGATGATCAGGCTGACTGGCTGGTGT 737  
DB |||||  
QY 919 GCTTAAACTGTTCAACAGCTTCTTTTGGAAACATTCAAATTCACAAAGTCAACTTCAGAT 978  
DB |||||  
QY 738 ACTAAAGCTCTTCAACAGCTTCTTCTGGAACATTCAGATTCAGAGGTCACCTGAGAT 797  
DB |||||  
QY 979 GGTAAAGCTGCAACTGAGACGAAATTTGCCGCTTCTGTTTCTACCAAGTTTCATAGATCCCCA 1038  
DB |||||  
QY 798 GGTAAAGCTGCAACTGAGACGAAATCTGCGCTCTTGTGTTTCTGCCGTTGCACAGATCCCCA 857  
DB |||||  
QY 1039 TATTGACTATGTGCTGCTCATCTTTCATCTTCTGCGCAATACATCAAAAGCACCATACAT 1098  
DB |||||  
QY 858 CATCGACTACTGTGCTGCTCACCCTTCATCTCTTCTGCCCAACATCAAAAGCTCCATACAT 917  
DB |||||  
QY 1099 TSCCTTCAGGCAATTAATCTCAACATCCCAATCTTCAGTACTCTTGCATCAATGCTGCGGG 1158  
DB |||||  
QY 918 CGCTTCGGGCAACAACTCAACATCCCATCTTCAGTACTCTTGCATCAATGCTGCGGG 977  
DB |||||  
QY 1159 CTTCTTTCAACGACGAAGGCTCGATGAAACACACAGATGGACGGAAGATGTTCTCTATAG 1218  
DB |||||  
QY 978 CTTTTTCAAGAGCGGAGGCTTGACGAAACTCCAGATGGACGCAAGACATTCGTACAG 1037  
DB |||||  
QY 1219 AGCTTTGCTCCATGGGCATATAGTTGAATTAATCTTCAGACAGCAGCAATTCCTTGAGATCTT 1278  
DB |||||  
QY 1038 AGCGTTGCTCCATGGGCATATAGTTGAATCTCTCCAGACAGCAGATTCCTTGAGATCTT 1097  
DB |||||  
QY 1279 CTTGGAAGGCACACGTTCTAGGAGTGGAAGAAACCTCTTGTGCTCGGSCAGGACTTTTGT 1338  
DB |||||  
QY 1098 CTTGGAAGGCACCCGCTCCCGAGTGGAAGACCTCTGTCGCCGCGGGCTCTCTGTC 1157  
DB |||||  
QY 1339 AGTTGTGGTAGATCTCTCTTACCAATGTCATCCAGACATCTTGTATATACCTGTGG 1398  
DB |||||  
QY 1158 AGTTGTAGTGGATCTCTGTCATCCCAACCATCCCTGACATCTCTGTCATCCCTGTGG 1217  
DB |||||  
QY 1399 AATCTCTATGATCGCAATTAATCGAGGTCACTACAATGTTGAACAACTCGGCAAACTTA 1458  
DB |||||  
QY 1218 CATCTCGTATGATCGGATTAATCGAGGTCACTACAATGTTGAACAGTTCGGCAAGCCAA 1277  
DB |||||  
QY 1459 GAAGATGAGCCTGTGGAGTGTGAAGAGGTGTTATTAGAATGTTTACGAAAAAACA 1518  
DB |||||  
QY 1278 GAAGATGAAGTCTCTCGAGTGTGGCAAGAGCGCTTATCAGAACTGCTCGGAAAAACA 1337  
DB |||||  
QY 1519 TGGTTGTGTCGAGTGGATTTTGCACAGCCATTTTCTTAAAGGAATATTTAGAAAGCCA 1578  
DB |||||

Db 1338 CGGCTATGTCAGTGAGCTTTGGCAGCCATTTTCTTTGAAGGAATATTAGAAAGGCCA 1397  
Qy 1579 AAGTCAGAAACCGGTGTCGTCTACTTTCCCTCGAGCAAGGGTTGTTACAGACTATACT 1638  
Db 1398 AAGTCAGAAACCTGTATCTGCTCCCTCTCTTTGGAGCAGACATGTTACAGCAATCT 1457  
Qy 1639 TCCTTCAAGACCCAGTGATGCTGCTGATGAAGGTAGAGACAGTCCTCAATTAATGATCCAG 1698  
Db 1458 TCCTTCAAGACCTGTATGCTGCTGCTGCCGAACATGAAGACATGTCAGTAATGATCGAG 1517  
Qy 1699 AATGCAACAGATGAATCCCTACGAGAGAGGTGATTGCAATCTGGCTGAGCATATCT 1758  
Db 1518 AAACCGGCAGACGAAGCTTCGGAAGGAGGCTGATCGCAACCTGGCGGAGCATTCT 1577  
Qy 1759 ATTCACTGTAGCAAGTCTGTGCCATTATGTCCACACACATTTGGCTTGCTCTCTCT 1818  
Db 1578 CTTACCGCCAGCAGTCTCGGCTATCATGTGCCACCCACATTTGGCTGCTCTCTCT 1637  
Qy 1819 CTACAGACACAGCGAGGAATGATCTCTCCAATTTGGTCGAAGACTTCTTTGTGATGAA 1878  
Db 1638 CTACAGACACAGCGAGGAATCCACTCTCCACGCTGTGTGAAGACTTCTTTGTGATGAA 1697  
Qy 1879 AGAGGAAGTCTGGCTCGTGATTTGACCTGGGTTCTCAGGAATTCAGAAGATGTAGT 1938  
Db 1698 GGAGGAAGTCTAGCTCGGATTTTGACCTGGGCTTCTCCGGGAATTCAGAAGATGTAGT 1757  
Qy 1939 AATGATGTCATACAGCTGCTGGGAATTTGTCTCAATTCACCCACACTAGCAGGAACGA 1998  
Db 1758 CATGATGCTATTAGCTTCTGGGAATGTTGTCTCAATTCACCCACACTAGCAGGAACGA 1817  
Qy 1999 TGAGTTTTTATCACCCCCAGACAACTGTCCCATCAGTCTTCGAACCTCACTCTACAG 2058  
Db 1818 TGAATTTCTTATCTCCAGCACAACTGTCCCGTCCGCTTTTGAATCACTCTACAG 1877  
Qy 2059 CAATGGGTACTTATGCTTTTATCATGAGGCGCATCATAGCTTCGAGCTTTATGCACT 2118  
Db 1878 CAATGGGTACTTATGCTTTTATCATGAGGCGCATCATAGCTTCGAGCTTTATGCACT 1937  
Qy 2119 TCTGAACAGAGGAGGACTGGGGGTCCTAGCAGCCACCTACCTACCTGATCAGCAGGA 2178  
Db 1938 CAGAAATAGAGGGGTTCGGAGGGTCTGCGAGGGCTTGGCAACCTGATCAGCAGGA 1997  
Qy 2179 GCAGCTGTGCGGAAGGGCGGAGCTGTGCTTACCTTTCTCCAATGAAGCAGCCTATCTC 2238  
Db 1998 GCAGCTGTGCGGAAGGGCGGAGCTGTGCTTACCTTTCTCTATGAAGTACCATTTTC 2057  
Qy 2239 ACTGCTTGCAGACATTTTACCAAGTCTGCCAATGAAGCAGTGAAGAAAGTTATCCAGTA 2298  
Db 2058 TCTGCCCTGCCAGACATTTTACAGGTTTGTCAAGAGACAGTAGGAAAGTTTATCCAGTA 2117  
Qy 2299 TGGCAATCTTACAGTGGCAGACGATGACCCAGGAAGATATCAGTCTTAGTCTTGCTGA 2358  
Db 2118 CGGAATTTCTCAGTGGCAGACAGATGACAGGAAGATGTGAGTCTGCGCTTGCGAGA 2177  
Qy 2359 GCAGAGTGGGCAAGAAGACTTCCAGAACCTTTGTCTTGGAGAGTGATGAAGAGATGA 2418  
Db 2178 GCAGAGTGGGAACAGAGACTTCCGAGGCTCTGAACCTGGAGAGTGACGAAGAGATGA 2237  
Qy 2419 AGACAGTGAATTTGGGGAGGAACAGCGGATTTGCTTACCTGAAGGTGAGCCCAATCCAAGGA 2478  
Db 2238 GCAGAGTGAATTTGGTGAAGGACGCTGATTTGCTTACCTGAAGGTGAGCCAGGCAAGGA 2297  
Qy 2479 GCACAGAGTTTATCACCTTCTTACAGACTCTTGGGCTTTGCTGGAGGCTTACAG 2538  
Db 2298 GCACAGCAATTTATCATCTTCTCAGAGGCTTCTGGGGCCCTCTGCTAGAGGCTACAG 2357  
Qy 2539 CTCTGCTGCATCTTTTGTTCACAACTTCAGTGGTCTCTGTTCAGAACCTGAGTATCTGCA 2598  
Db 2358 CTCTGCTGCATCTTTTGTTCACAACTTTCGGGGCCAGTTCGGAGTCTGAGTACCTGCA 2417  
Qy 2599 AAGTTGCACAATACTTAATAACAGGAACAGAAAGATTTGTCAGTATATGCTGAGAG 2658

Db 2418 GAAGCTGCACAGGTACCTTTCTCACAGGACGGAGAGGAACGTCGCGGTGTACGCTGAGAG 2477  
Qy 2659 TGCCACATATTGCTTGTGAAGAATGCTGTCAAAATGTTTAAGGATATTGGGGTTTTTCAA 2718  
Db 2478 TGCCACATATTGCTTGTGAAGAATGCTGTCAAAATGTTTAAGGACATCGGGGTTTTCAA 2537  
Qy 2719 GGAGACCAAAACAAAGAGAGTGTCTGTTTATAGAACTGAGCAGCACTTTTCTACCTCAATG 2778  
Db 2538 AGAGACCAAGCAGAAGCGAGCTGTCTTATAGAACTGAGCAACCACTTTCTACCTCAGGG 2597  
Qy 2779 CAACCGCAAAAACCTTCTAGAAATATATCTGAGTTTGTGTGCTGTAG 2827  
Db 2598 CAGCCGCGAGAAGCTCCTGGAATATCATTTCTGAGCTTCGTGTGTGTAG 2646

RESULT 6  
US-09-833-381-1961  
; Sequence 1961, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1: Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1961  
; LENGTH: 612  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-833-381-1961

Query Match 19.4%; Score 581.6; DB 10; Length 612;  
Best Local Similarity 99.0%; Pred. No. 1.4e-176;  
Matches 606; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy 1401 TCTCTATGATCGCATTTATCGAAGTCACTCAATGTTGAACAACTGGGCAAACTAAGA 1460  
Db 1 TCCGCTATGATCGCATTTATCGAAGTCACTCAATGTTGAACAACTGGGCAAACTAAGA 60  
Qy 1461 AGAATGAGAGCCTCTGAGTGTAGCAAGAGGTGTTATTAGAACTGTACGAAAAAATATG 1520  
Db 61 AGAATGAGAGCCTCTGAGTGTAGCAAGAGGTGTTATTAGAACTGTACGAAAAAATATG 120  
Qy 1521 GTTGTGTCGAGTGGATTTTGCACAGCCATTTTCTTAAAGGAATATTAGAAAGCCAAA 1580  
Db 121 GTTGTGTCGAGTGGATTTTGCACAGCCATTTTCTTAAAGGAATATTAGAAAGCCAAA 180  
Qy 1581 GTCAAAA - CCGGTGTCTGCTTACTTTCCCTGAGCAAGGTTGTTACAGCTATACTT 1639  
Db 181 GTCAAAAACCGGTGTCTGCTTACTTTCCCTGAGCAAGGTTGTTACAGCTATACTT 240  
Qy 1640 CTTTCAAGACCCAGTGTGCTGATGAAGGTAGACACGTCCTCAATTAATGAGTCCAGA 1699  
Db 241 CTTTCAAGACCCAGTGTGCTGATGAAGGTAGACACGTCCTCAATTAATGAGTCCAGA 300  
Qy 1700 AATCAACACAGATGAATCCCTACGAAGGAGGTGTTGATTCAAATCTGGCTGAGCATATTCTA 1759  
Db 301 AATCAACACAGATGAATCCCTACGAAGGAGGTGTTGATTCAAATCTGGCTGAGCATATTCTA 360  
Qy 1760 TTCACTGTAGCAAGTCTGTGCCATTTATGTCACACATGTTGGTGTGCTGCTCTC 1819  
Db 361 TTCACTGTAGCAAGTCTGTGCCATTTATGTCACACACATGTTGGTGTGCTGCTCTC 420  
Qy 1820 TACAGACACAGCGAGGAATGATCTCTCAATGTTGGTTCGAAGACTTCTTTGTGATGAAA 1879  
Db 421 TACAGACACAGCGAGGAATGATCTCTCAATGTTGGTTCGAAGACTTCTTTGTGATGAAA 480  
Qy 1880 GAGGAAGTCTTGGCTCGTGTGATTTTGAAGTCTTCTCAGGAAATTCAGAAAGATGT - AGT 1938



```

; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7822
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-7822

Query Match      12.8%; Score 385.4; DB 11; Length 399;
Best Local Similarity 99.5%; Pred. No. 3.2e-113;
Matches 397; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 630 CTCACACAGACACCGCGATGGCTTGCAGAGCGCTTCTTACGCTCTTTTATTCAAG 689
Db 1 CTCACACAGACACCGCGATGGCTTGCAGAGCGCTTCTTACGCTCTTTTATTCAAG 60

QY 690 AGCGAGATGTGATGAAGGCGCATGTTGCCACCAATGTGACTGAAATGTGC-TGAACAGC 748
Db 61 AGCGAGATGTGATGAAGGCGCATGTTGCCACCAATGTGACTGAAATGTGCTTGNACAGC 120

QY 749 AGTAGAGTACAGAGCGCAATTGCAGAGTGGCTGTGAAATTAACCCCTGATGGTTCTGCC 808
Db 121 AGTAGAGTACAGAGCGCAATTGCAGAGTGGCTGTGAAATTAACCCCTGATGGTTCTGCC 180

QY 809 CAGCAGCAATCAAAAGCGGTTTAAACAAAGTGAAGAAAGCAAGCTTAAAGATTTCTCAAGAA 868
Db 181 CAGCAGCAATCAAAAGCGGTTTAAACAAAGTGAAGAAAGCAAGCTTAAAGATTTCTCAAGAA 240

QY 869 ATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGACTGGGTGGTCTGTCTAAAGCTG 928
Db 241 ATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGACTGGGTGGTCTGTCTAAAGCTG 300

QY 929 TTCAACAGCTCTTTTGGAAACATTCAAATTCACAAAGTCAACTTGAGATGGTTAAAGCT 988
Db 301 TTCAACAGCTCTTTTGGAAACATTCAAATTCACAAAGTCAACTTGAGATGGTTAAAGCT 360

QY 989 GCAACTGACAGCAATTTCCCGCTCTGTTTCTACCAAGTT 1027
Db 361 GCAACTGACAGCAATTTCCCGCTCTGTTTCTACCAAGTT 399

RESULT 10
; Sequence 1964, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1959
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1959

Query Match      11.5%; Score 346.6; DB 10; Length 396;
Best Local Similarity 95.9%; Pred. No. 1.1e-100;
Matches 375; Conservative 0; Mismatches 13; Indels 3; Gaps 2;

QY 2178 AGCAGCTGGTGGGAGGCGCGCAGCCTGTGCTACCTTCTCTCCAATGAAGGCAACCATCT 2237
Db 1 AGCAGCTGGTGGGAGGCGCGCAGCCTGTGCTACCTTCTCTCCAATGAAGGCAACCATCT 60

QY 2238 CACTGCTTGCAGACATTTTACCAAGTCTGCCATGAAGAGTATGAGTATTCAGT 2297
Db 61 CACTGCTTGCAGACATTTTACCAAGTCTGCCATGAAGAGTATGAGTATTCAGT 120

QY 2298 ATGGCATTTCTACAGTGGCAGGACCATGATCAGGAGAGATATCAGTCTTAGTCTTCTG 2357
Db 121 ATGGCATTTCTACAGTGGCAGGACCATGATCAGGAGAGATATCAGTCTTAGTCTTCTG 180

QY 2358 AGCAGCAGTGGGCAAGAAGCTTCCAGAACCTTTGTCTTGGAGAGTATGATGAAGAGATG 2417
Db 181 AGCAGCAGTGGGCAAGAAGCTTCCAGAACCTTTGTCTTGGAGAGTATGATGAAGAGATG 240

QY 2418 AAGACAGTACTTTGGGGAGGAAACAGCAGATGCTACTGAA--GGTGAAGCCAAATCAA 2475
Db 241 AAGACAGTACTTTGGGGAGGAAACAGCAGATGCTACTGAAAGGTGAGCCCAATCAA 300

; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7822
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-7822

Query Match      11.6%; Score 348.4; DB 10; Length 440;
Best Local Similarity 98.3%; Pred. No. 3.2e-101;
Matches 349; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7822
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-7822

Query Match      12.8%; Score 385.4; DB 11; Length 399;
Best Local Similarity 99.5%; Pred. No. 3.2e-113;
Matches 397; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 630 CTCACACAGACACCGCGATGGCTTGCAGAGCGCTTCTTACGCTCTTTTATTCAAG 689
Db 1 CTCACACAGACACCGCGATGGCTTGCAGAGCGCTTCTTACGCTCTTTTATTCAAG 60

QY 690 AGCGAGATGTGATGAAGGCGCATGTTGCCACCAATGTGACTGAAATGTGC-TGAACAGC 748
Db 61 AGCGAGATGTGATGAAGGCGCATGTTGCCACCAATGTGACTGAAATGTGCTTGNACAGC 120

QY 749 AGTAGAGTACAGAGCGCAATTGCAGAGTGGCTGTGAAATTAACCCCTGATGGTTCTGCC 808
Db 121 AGTAGAGTACAGAGCGCAATTGCAGAGTGGCTGTGAAATTAACCCCTGATGGTTCTGCC 180

QY 809 CAGCAGCAATCAAAAGCGGTTTAAACAAAGTGAAGAAAGCAAGCTTAAAGATTTCTCAAGAA 868
Db 181 CAGCAGCAATCAAAAGCGGTTTAAACAAAGTGAAGAAAGCAAGCTTAAAGATTTCTCAAGAA 240

QY 869 ATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGACTGGGTGGTCTGTCTAAAGCTG 928
Db 241 ATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGACTGGGTGGTCTGTCTAAAGCTG 300

QY 929 TTCAACAGCTCTTTTGGAAACATTCAAATTCACAAAGTCAACTTGAGATGGTTAAAGCT 988
Db 301 TTCAACAGCTCTTTTGGAAACATTCAAATTCACAAAGTCAACTTGAGATGGTTAAAGCT 360

QY 989 GCAACTGACAGCAATTTCCCGCTCTGTTTCTACCAAGTT 1027
Db 361 GCAACTGACAGCAATTTCCCGCTCTGTTTCTACCAAGTT 399

RESULT 10
; Sequence 1964, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1964
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(440)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1964
```

RESULT 13  
US-09-960-352-8821  
; Sequence 8821, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Bvatt, John C.

RESULT 14  
 US-09-960-352-742/c  
 ; Sequence 742, Application US/09960352  
 ; Patent No. US20020137139A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Warren, Wesley C.  
 ; APPLICANT: Tao, Nengbing  
 ; APPLICANT: Byatt, John C.  
 ; APPLICANT: Mathialagan, Nagappan  
 ; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
 ; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION  
 ; FILE REFERENCE: 16511.006/37-21 (10298)C  
 ; CURRENT APPLICATION NUMBER: US/09/960,352  
 ; CURRENT FILING DATE: 2001-09-24  
 ; NUMBER OF SEQ ID NOS: 15112  
 ; SEQ ID NO 742  
 ; LENGTH: 418  
 ; TYPE: DNA  
 ; ORGANISM: Bos taurus  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (395)  
 ; OTHER INFORMATION: unsure at all n locations  
 ; OTHER INFORMATION: Clone ID: 04-LIB2809-001-Q1-E1-A11  
 US-09-960-352-742



```

RESULT 15
US-09-814-353-16182/c
; Sequence 16182, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lille, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16182
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-16182

```

Query Match 8.9%; Score 268.6; DB 13; Length 379;  
Best Local Similarity 96.6%; Pred. No. 1.9e-75;

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 22:59:07 ; Search time 4035.2 Seconds  
(without alignments)  
18087.435 Million cell updates/sec

Title: US-09-935-290-1  
Perfect score: 3003  
Sequence: 1 ttggccaccaggctgctgcyg.....tctgagatcttccagcttt 3003

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estma:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est6:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1898.2	63.2	3305	11 AK047295	AK047295 Mus muscu
2	889.2	29.6	1043	12 BQ062851	BQ062851 AGENCOURT
3	856.6	28.6	1055	12 BQ057957	BQ057957 AGENCOURT
4	848.2	28.2	918	13 BQ437820	BQ437820 AGENCOURT

5	794.8	26.5	968	12 BQ058377	BQ058377 AGENCOURT
6	703.8	23.4	965	10 BF794257	BF794257 602255455
7	684	22.8	687	9 AL598680	AL598680 DKF2p313P
8	582.4	19.4	776	14 CD000106	CD000106 AGENCOURT
9	573	19.1	573	14 CB269030	CB269030 1007937 H
10	564.2	18.8	637	14 CB268766	CB268766 1007672 H
11	561.4	18.7	753	13 BU703927	BU703927 UI-M-FOO-
12	555.6	18.5	799	10 BG704442	BG704442 602688632
13	547.2	18.2	580	10 AW976326	AW976326 EST388435
14	530.8	17.7	602	9 AL603562	AL603562 DKF2p686P
15	530	17.6	988	10 BG034520	BG034520 602302638
16	529.2	17.6	1137	12 BM471357	BM471357 AGENCOURT
17	497	16.6	498	13 BQ694826	BQ694826 1000969 H
18	497	16.6	553	9 AL596527	AL596527 DKF2p451D
19	496.8	16.5	550	13 BQ694822	BQ694822 1000965 H
20	490.4	16.3	492	14 CB269165	CB269165 1008072 H
21	488.6	16.3	587	14 CA406237	CA406237 1002386 H
22	487.6	16.2	690	12 BI691948	BI691948 603311882
23	483.6	16.1	1037	10 BF037399	BF037399 601460764
c					AA769055 oa74a10.s
24	480	16.0	656	9 AA769055	AA769055 oa74a10.s
25	467.8	15.6	825	13 BU263739	BU263739 603815507
26	446	14.9	455	13 BQ694370	BQ694370 1000513 H
27	445	14.8	589	14 CA889049	CA889049 B0150D03-
28	440.6	14.7	746	14 BY741039	BY741039 BX741039
29	420.4	14.0	859	12 BI738244	BI738244 603361683
c					BM414848 ECC00320
30	398	13.3	446	12 BM414848	BM414848 ECC00320
31	395.8	13.2	469	12 BI551617	BI551617 603194269
32	393	13.1	564	9 AA277375	AA277375 va81g09.r
33	388.8	12.9	646	12 BI100769	BI100769 602886066
34	378.4	12.6	535	14 CA896337	CA896337 B0198D07-
35	369.6	12.3	527	14 CD554026	CD554026 B0370D05-
36	368.6	12.3	400	14 R73257	R73257 y109g05.r1
37	366.6	12.2	519	14 CA887606	CA887606 B0138C12-
38	362.8	12.1	491	14 CA885724	CA885724 B0122A11-
39	343.4	11.4	347	14 CB265571	CB265571 1004476 H
c					AA172653 ms19c01.r
40	340.2	11.3	571	9 AA172653	AA172653 ms19c01.r
41	329.8	11.0	411	12 BM146906	BM146906 TCAAP1E75
42	324	10.8	419	12 BM710854	BM710854 UI-E-DX0-
43	321.8	10.7	823	14 CD362496	CD362496 AGENCOURT
c					CD502509 CDA57-B03
44	321.4	10.7	1310	14 CD502509	CD502509 CDA57-B03
45	318	10.6	713	13 BU444448	BU444448 603766560

ALIGNMENTS

RESULT 1  
AK047295  
LOCUS  
DEFINITION  
AK047295 3305 bp mRNA linear HTC 05-DEC-2002  
Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length  
enriched library, clone:B930046K04 product:glycerol-3-phosphate  
acyltransferase, mitochondrial, full insert sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED

AK047295.1 GI:26092071  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
93279253  
10349636  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

REFERENCE  
AUTHORS

3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsui, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

REFERENCE  
AUTHORS

4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,  
Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,  
Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.  
and Hayashizaki, Y.

## TITLE

Functional annotation of a full-length mouse cDNA collection

JOURNAL  
MEDLINE  
PUBMED

Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

REFERENCE  
AUTHORS

5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.

## TITLE

Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs

JOURNAL  
MEDLINE  
PUBMED

Nature 420, 563-573 (2002)  
6 (bases 1 to 3305)

REFERENCE  
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.

## TITLE

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues  
Please visit our web site for further details.

## JOURNAL

http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

## FEATURES

Location/Qualifiers

source	1..3305	/organism="Mus musculus"	
		/mol_type="mRNA"	
		/strain="CS7BL/6J"	
		/db_xref="PANTOM_DB:B930046K04"	
		/clone_xref="taxon:10090"	
		/clone="B930046K04"	
		/tissue_type="cerebellum"	
		/clone_lib="RIKEN full-length enriched mouse cDNA library"	
		/dev_stage="10 days neonate"	
misc_feature	156..2835	/note="glycerol-3-phosphate acyltransferase, mitochondrial (MGDI) [GI:109162, GB NM_008149, evidence: BLASTN, 99%, match=2883]"	
		putative"	
polyA_signal	3281..3286	/note="putative"	
	3305	/note="putative"	
polyA_site			
BASE COUNT	819 a 861 c 832 g 793 t		
ORIGIN			
Query Match	63.2%;	Score 1898.2;	DB 11; Length 3305;
Best Local Similarity	82.9%;	Pred. No. 0;	
Matches 2229;	Conservative	0; Mismatches 445;	Indels 15; Gaps 5;
QY	312	CACATGATTGGGAATTACACTTTGTGACATGGATGAATCTGCACCTGACCTTGGTACAA	371
DB	127	CACATGTTGGGACTTGGACGTTCTGCCATGGAGGAGTCTTCAGTGACAGTTGGCACA	186
QY	372	TAGATGTTTCTTATCTGCCACATTTTCATCAGAATACAGTGTGGTGCATGAAGCACACA	431
DB	187	TAGAGCTTTCTTATCTGCCACGTTTCATCGGAATACAGCCTTGGCGGATGAACACACA	246
QY	432	GTGAGGATGGGGTGAGTGTGGCTTTAGACCCACCGCTTTCAGATCTGCAACTTTAAAT	491
DB	247	GTGAGGACTGGGTGTGACTGTGGGTTTCAACCTTACCTTCTTCAGATCTGCAACTGAAT	306
QY	492	GGAAAGAAAGCCTAATGAGTCGGAAGGCAATTTGTTGGAAGATGTTCTTACTCTCGCA	551
DB	307	GGAGGAGAGCCTTATGAGCCGGAAGAGGCCCTTCTGTTGGAAGGCTGCTCTTCTCTGCA	366
QY	552	CTCCCGACAGCTGGGACAAATTTTCAACCCAGATATCCCGTCTTTGGGTTTGGCGAATG	611
DB	367	CGCCACAGAGCTGGGAAGGTTTTCACCCAGATATCCATCTCTGGGTTTGGCGAATG	426
QY	612	TTATTATATCAATGAATCTACAGACACCGCGGATGGCTTTCAGACAGCGCTTTCTT	671
DB	427	TTATTATATCAATGAATGAACGACACAGAGGATGGCTGGCGAGACGGCTGTCTT	486
QY	672	ACGTTCTTTTATCAAGAGCGAGATGTGCATTAAGGCGATGTTTGCACCAATGTGACTG	731
DB	487	ACATCTCTTTTGTTCAGAGCGAGAGCTCCATTAAGGCGATGTTTGCACCAAGTGTACTG	546
QY	732	AAATGTGCTGAACAGCAGTAGAGTACAGAGGCAATTTGCAAGTGGCTGCTGAATTA	791
DB	547	AGATGTACTAGCAGCAGCAGTCCAGAGGCAATTTGCTGAGTGGCTGGGAGTTGA	606
QY	792	ACCTGATGGTCTGCGCCAGCAGCAATCAAAAGCGGTTAAAGAGTGAAGAAAGCTA	851
DB	607	ACCCAGATGGATCTGCCAGCAGCAGTCCAAAGCCATCCAGAGGTTGAAGAAAGGCA	666
QY	852	AAAGGATTTTCAAGAAATGTTGCCATGCTCTCACCGGCAATGATCAGACTGACCTGGGT	911
DB	667	GGAGATCTCTCAGGAGATGGTGGCCACCGTCTCCCGAGGATGATCAGGCTGACCTGGCT	726
QY	912	GGGTGCTGCTAAAACCTGTTCAACAGCTTCTTTTGGAAACATTCAAATTCACAAAGGTCAAC	971
DB	727	GGGTGTTACTAAGCTTTCACAGCTTCTTCTGGAACATTCAGATTCAAGAGGTCAAC	786
QY	972	TTGAGATGTTAAAGCTGCACTGAGACGAATTTGCCGCTTCTGTTTCTACAGTTCAATA	1031
DB	787	TCGAGATGTTCAAGGCTGCACTGAGACGAACCTGCGCTCTTGTCTTCTGCGGTGCACA	846



REFERENCE 1 (bases 1 to 1043)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: Lou Staudt  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM2094 row: 9 column: 16  
 High quality sequence stop: 608.  
 Location/Qualifiers  
 1..1043  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5924391"  
 /tissue\_type="lymphoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_99"  
 /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."  
 BASE COUNT 296 a 238 c 231 g 278 t  
 ORIGIN  
 Query Match 29.6%; Score 889.2; DB 12; Length 1043;  
 Best Local Similarity 95.7%; Pred. No. 1.1e-241;  
 Matches 957; Conservative 0; Mismatches 38; Indels 5; Gaps 4;  
 QY 212 AGGACAGCGGCTCCCTGTTGATGGACATTTCTGCACCGGAAACTGATAGCTGAGTCC 271  
 DB 43 AGGACAGCGGCTCCCTGTTGATGGACATTTCTGCACCGGAAACTGATAGCTGAGTCC 102  
 QY 272 TGAAGTTTTATGATCAACAGAGAACCTTTCATCCAGACATGATTTGGGAATTACA 331  
 DB 103 TGAAGTTTTATGATCAACAGAGAACCTTTCATCCAGACATGATTTGGGAATTACA 162  
 QY 332 CTTTGTGACATGATGAATCTGCACCTGACCTGGTACAATAGATGTTCTTATCTGCCA 391  
 DB 163 CTTTGTGACATGATGAATCTGCACCTGACCTGGTACAATAGATGTTCTTATCTGCCA 222  
 QY 392 CATTTCATCAGAAATACAGTGTGGTGCATGTAAGACACAAAGTGAAGGAATGGGTGAGTGT 451  
 DB 223 CATTTCATCAGAAATACAGTGTGGTGCATGTAAGACACAAAGTGAAGGAATGGGTGAGTGT 282  
 QY 452 GGCTTTAGACCCCGCTTTCAGATCTGCACTTAAATGGAAGGAAGCCTTAATGAGT 511  
 DB 283 GGCTTTAGACCCCGCTTTCAGATCTGCACTTAAATGGAAGGAAGCCTTAATGAGT 342  
 QY 512 CGGAAAAGGCGCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAA 571  
 DB 343 CGGAAAAGGCGCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAA 402  
 QY 572 TTTTTCACCCAGATATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAGT 631  
 DB 403 TTTTTCACCCAGATATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAGT 462  
 QY 632 CACACAGACACCGCGATGGCTTGCAGAGCGCTTCTTACGTTCTTTTATTCAGAG 691  
 DB 463 CACACAGACACCGCGATGGCTTGCAGAGCGCTTCTTACGTTCTTTTATTCAGAG 522  
 QY 692 CGAGATGTGCATAAGGCGATGTTTGCCACCAATGTGACTGAAAATGTGCTGAACAGCAGT 751

Db 523 CGAGATGTGCATAAGGCGATGTTTGCCACCAATGTGACTGGAAATGTGCTGAACAGCAGT 582  
 QY 752 AGAGTACAGAGGCAATTCAGAGTGGCTGCTGAATTAACCCCTGATGTTTCTGCCAG 811  
 Db 583 AGAGTACAGAGGCAATTCAGAGTGGCTGCTGAATTAACCCCTGATGTTTCTGCTCAG 642  
 QY 812 CAGCAATCAAAAGCGCTTAAACAAAGTGAAGAAAGCTTAAAGGATTTCTTCAAGAAATG 871  
 Db 643 CAGCAATCAAAAGCGCTTAAACAAAGTGAAGAAAGCTTAAAGGATTTCTTCAAGATATG 702  
 QY 872 GTTGCCACTGTCTCACCGGCAATGATCAGACTGAGTGGTGGTGGTGTCTTAAACCTGTTTC 931  
 Db 703 GTTGCCACTGTCTCACCGGCAATGATCAGACTGAGTGGTGGTGGTGTCTTAAACCTGTTTC 762  
 QY 932 AACAGCTTCTTTTGGAAACATTCAAATTCACAAGGTCAA-CATTGAGATGTTTAAAGCTGC 990  
 Db 763 AACAGCTTCTTTTGGAAACATTCAAATTCACAAGGTCAAACCTTTGAGATGTTTAAAGCTGC 822  
 QY 991 AACTGAGACGAATTTGCCGCTTCTGTTTCTACAGTTCATAGATCCCATATGACTATCT 1050  
 Db 823 CACTGAGACGAATTTGCCGCTTCTGTTTCTACAGTTCATAGATCCCATATGACTATCT 882  
 QY 1051 -GCTGCTCACTTTCATTTCTTCTGCCATPAACATCAAGCACATACATGTTCTCAGGCA 1109  
 Db 883 GGCTGCTCACTTTCATTTCTTCTGCCATPAACATCAAGCACATACATGTTCTCAGGCA 942  
 QY 1110 ATATCTCAACATCCCAATCTTCAGTACCTTCATCATAAGCTT--GGGGGCTTCTTCAT 1167  
 Db 943 AATATCTCAACATCCCAATCTTCAGTACCTTCATCATAAGCTTGGGGGGGCTTCTTCT 1002  
 QY 1168 ACACCAAGGCTCGA-TGAACACACAGATGGACGGAAGA 1206  
 Db 1003 TCAACAAAGCTTCATTGAAACCCCAATGGACGGTAAAA 1042

RESULT 3  
 LOCUS BQ057957  
 DEFINITION AGENCOURT\_6924421 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5813732  
 5', mRNA sequence.  
 ACCESSION BQ057957  
 VERSION BQ057957.1 GI:19817297  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1055)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: Lou Staudt  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM2066 row: d column: 21  
 High quality sequence stop: 666.  
 Location/Qualifiers  
 1..1055  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5813732"  
 /tissue\_type="lymphoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_99"  
 /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:  
 FEATURES  
 source







```
QY 873 TTGGCACTGCTCTCACGGCAATGATCAGACTGCTGGGTGGGTGCTGCTGCTAAACCTGTTCA 932
Db 702 TTGGCACTGCTCTCACGGCAATGATCAGACTGCTGGGTGGGTGCTGCTGCTAAACCTGTTCA 761
QY 933 ACAGCTTCTTTTGGAAACATTCAAAATTCACAAAGGTCACCTTCAGATGTTTAAAGCTGCAA 992
Db 762 ACAGCTTCTTTTGGAAACATTCAAAATTCACAAAGGTCACCTTCAGATGTTTAAAGCTGCAA 821
QY 993 CT--GAGACGAATTGGCGCTTCTGTTTCTACCAAGTT--CATAGATCCCATATTGACTA 1047
Db 822 CTGGAGACGAATTTGGCGCTTCTGTTTCTACCAAGTTTCATAGATCCCATATTGACTA 881
QY 1048 T-CTGCTGCTAC-TTTCATCTCTCTTCCCATTAAC 1081
Db 882 TCCTGCTGCTAC-TTTCATCTCTCTTCCCATTAAC 917

RESULT 6
BF794257
LOCUS
DEFINITION
602255455F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4338997 5',
mRNA sequence.
ACCESSION
BF794257
VERSION
BF794257.1 GI:12099311
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 965)
NIH-MGC http://mgi.mcni.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LILAM9948 row: m column: 14
High quality sequence stop: 696.
Location/Qualifiers
1..965
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4338997"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 273 a 222 c 226 g 244 t
ORIGIN

Query Match 23.4%; Score 703.8; DB 10; Length 965;
Best Local Similarity 97.6%; Pred. No. 8.6e-189;
Matches 758; Conservative 0; Mismatches 12; Indels 7; Gaps 4;

QY 1304 GGAAAACTCTTGCTCGGCAGAGACTTTTGTCAGTTGGTATGATCTGCTAC 1363
Db 1 GGAAAACTCTTGCTCGGCAGAGACTTTTGTCAGTTGGTATGATCTGCTAC 60
QY 1364 ATGTGTCATCCAGACATCTTGATATACCTGTTGAATCTCTATGATCGCATTCGAA 1423
Db 61 AATGTGTCATCCAGACATCTTGATATACCTGTTGAATCTCTATGATCGCATTCGAA 120
```

```
QY 1424 GGTCACTACAATGTGTGAACAACTGGGCAAACTAAGAGAAATGAGAGCTCTGGAGTGA 1483
Db 121 GGTCACTACAATGTGTGAACAACTGGGCAAACTAAGAGAAATGAGAGCTCTGGAGTGA 180
QY 1484 GCAGAGGTGTATAGAAATGTTACGAAAAAATCTATGGTTGTGTCGAGTGGATTTGCA 1543
Db 181 GCAGAGGTGTATAGAAATGTTACGAAAAAATCTATGGTTGTGTCGAGTGGATTTGCA 240
QY 1544 CAGCCATTTCTTTAAAGGAATATTTAGAAAGCCAAAGTTCAGAAACCGGTGCTGCTCTGA 1603
Db 241 CAGCCATTTCTTTAAAGGAATATTTAGAAAGCCAAAGTTCAGAAACCGGTGCTGCTCTGA 300
QY 1604 CTTTCCCTGGAGCAAGCGTTGTTACCACTATATCTCTTCAAGACCCAGTGTGCTGCT 1663
Db 301 CTTTCCCTGGAGCAAGCGTTGTTACCACTATATCTCTTCAAGACCCAGTGTGCTGCT 360
QY 1664 GATGAAGGTAGAGACACGTCCTCAATTAAGTCCAGAAATGCAACAGATGAATCCCTACGA 1723
Db 361 GATGAAGGTAGAGACACGTCCTCAATTAAGTCCAGAAATGCAACAGATGAATCCCTACGA 420
QY 1724 AGGAGGTGATTGCAAAATCTGGCTGAGCATATTTCTTACTGCTAGCAAGTCTCTGTGCC 1783
Db 421 AGGAGGTGATTGCAAAATCTGGCTGAGCATATTTCTTACTGCTAGCAAGTCTCTGTGCC 480
QY 1784 ATTATGTCACACACATTTGTGGCTTGCTCTCTCTACAGACACAGCAGGGAATTGAT 1843
Db 481 ATTATGTCACACACATTTGTGGCTTGCTCTCTCTACAGACACAGCAGGGAATTGAT 540
QY 1844 CTCTCCACATTTGTCGAAGACTTCTTTGTGATGAAGAGGAAGTCTCTGGCTGCTGATTTT 1903
Db 541 CTCTCCACATTTGTCGAAGACTTCTTTGTGATGAAGAGGAAGTCTCTGGCTGCTGATTTT 600
QY 1904 GACCTGGGGTCTTCAGGAAATTCAGAAAGATCTAGTAATGTCATGCATACACAGTCTGCGGA 1963
Db 601 GACCTGGGGTCTTCAGGAAATTCAGAAAGATCTAGTAATGTCATGCATACACAGTCTGCGGA 660
QY 1964 -AATTGTGTCAATCAATCACCCACACTAGCAGGAACGATGAGTTTATATCAATCCAGCAGC 2022
Db 661 CAATTGTGTCAATCAATCACCCACACTAGCAGGAACGATGAGTTTATATCAATCCAGCAGC 718
QY 2023 AACTGTCCATCAGCTTTCGAACTCAACTTCTCAGCAATGGGGTACTTTCATGCTT 2079
Db 719 CACTGTCC--TCAGTCTTCGACTCAAGTTCTACAGGA--TGGGTACTTTCATGCTT 771

RESULT 7
AL598680
LOCUS
DEFINITION
DKFZp313P1521_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
DKFZp313P1521 5', mRNA sequence.
ACCESSION
AL598680
VERSION
AL598680.1 GI:15161371
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 687)
Poustka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann, S.).
EST (Poustka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann
, S.)
Unpublished
Contact: Poustka A.J.
Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innistrasse 73, 14195 Berlin, Germany
Tel.: +49-30-84131623
Fax: +49-30-84131128
Email: poustka@mping-berlin-dahlem.mpg.de
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
```

Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.  
No sl sequence available.  
This clone (DKFZp313P1521) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

Location/Qualifiers  
1..687  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp313P1521"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="313 (synonym: hlcc2)"  
/notes="vector: pTriplEx2; Site\_1: SfiIA; Site\_2: SfiIB; cDNA-collection"  
202 a 139 c 168 g 178 t

Query Match 22.8%; Score 684; DB 9; Length 687;  
Best Local Similarity 100.0%; Pred. NO. 3.2e-183;  
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2220 CCAATGAAGGCACCATCTCACTGCTTCCAGACATTTTACCAAGTCTGCCATGAACAG 2279  
Db 1 CCAATGAAGGCACCATCTCACTGCTTCCAGACATTTTACCAAGTCTGCCATGAACAG 60

Qy 2280 TAGAAGATTATCCAGTATGGCATTTCTTACAGTGGCAGACGATGACAGAGATTA 2339  
Db 61 TAGAAGATTATCCAGTATGGCATTTCTTACAGTGGCAGACGATGACAGAGATTA 120

Qy 2340 TCAGTCTCTGCTGTCGACGAGTGGCAGACGATGTCAGACCTTCTCTTGGGA 2399  
Db 121 TCAGTCTCTGCTGTCGACGAGTGGCAGACGATGTCAGACCTTCTCTTGGGA 180

Qy 2400 GAAGTGTATGAAGATGAAGACAGTCACTTTGGGAGGAGACGACGATGTCACCTGA 2459  
Db 181 GAAGTGTATGAAGATGAAGACAGTCACTTTGGGAGGAGACGACGATGTCACCTGA 240

Qy 2460 AGGTGAGCCAAATCAAGGACACGACGATTTATCACTTCTTACAGAGACTCTCTGGGC 2519  
Db 241 AGGTGAGCCAAATCAAGGACACGACGATTTATCACTTCTTACAGAGACTCTCTGGGC 300

Qy 2520 CTTTGTCTGAGGCTCAGCTCTGCTGCCATCTTGTTCACACTTCAGTGCCTGCTTC 2579  
Db 301 CTTTGTCTGAGGCTCAGCTCTGCTGCCATCTTGTTCACACTTCAGTGCCTGCTTC 360

Qy 2580 CAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTTAATACAGACAGAAAGAAATG 2639  
Db 361 CAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTTAATACAGACAGAAAGAAATG 420

Qy 2640 TTGCAGTATATGCTGAGAGTGCACATATTTGCTGTGAAGATGCTGTGAAAATGTTTA 2699  
Db 421 TTGCAGTATATGCTGAGAGTGCACATATTTGCTGTGAAGATGCTGTGAAAATGTTTA 480

Qy 2700 AGGATATTTGGGTTTTCAAGGAGACCAACAAAGAGAGTGTCTGTTTGAAGTGAAGCA 2759  
Db 481 AGGATATTTGGGTTTTCAAGGAGACCAACAAAGAGAGTGTCTGTTTGAAGTGAAGCA 540

Qy 2760 GCATTTTCTACCTCAATGCACCGACAAAACCTTCTAGAAATATATTCAGTGTGTTGG 2819  
Db 541 GCATTTTCTACCTCAATGCACCGACAAAACCTTCTAGAAATATATTCAGTGTGTTGG 600

Qy 2820 TGCTGTAGGTAACTGTGGCACTGCTGGCAATGAAGGTATGAGATGATTCCTGTGTAG 2879  
Db 601 TGCTGTAGGTAACTGTGGCACTGCTGGCAATGAAGGTATGAGATGATTCCTGTGTAG 660

Qy 2880 GTACCAGCTTCTGGCTCAAGGTT 2903  
Db 661 GTACCAGCTTCTGGCTCAAGGTT 684

## RESULT 8

CD000106

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CD000106 776 bp mRNA linear EST 01-MAY-2003  
AGENCOURT 13642874 NIH\_MGC\_186 Homo sapiens cDNA clone  
IMAGE:30322827 5', mRNA sequence.

CD000106

CD000106.1 GI:30294625

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 776)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDCM124 row: h column: 04

High quality sequence stop: 549.

Location/Qualifiers

source

1..776

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30322827"

/lab\_host="DH10B (T1 phage-resistant)"

/note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: SfiI

(ggcattatggcc); Site 2: SfiI (ggcgctcgcc); Library is

oligo-dT primed and directionally cloned. cDNA was

prepared from a pooled samples of tissues from skin,

meninges, duramater, pia mater and choroid plexus. 5'

and 3' adaptors were used in cloning as follows: 5'

adaptor sequence: 5'-CACGCCATATGGC-3' and 3' adaptor

sequence: 5'-ATTCTAGAGCGCGCGCCGACATG-TT(30)BN-3'

(where B = A, C, or G and N = A, C, G, or T). Average

insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies

contained inserts by PCR. This library was enriched for

full-length clones and was constructed by Clontech

Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC

Library"

BASE COUNT 208 a 180 c 199 g 189 t

ORIGIN

Query Match 19.4%; Score 582.4; DB 14; Length 776;

Best Local Similarity 95.1%; Pred. NO. 3.3e-154;

Matches 643; Conservative 0; Mismatches 28; Indels 5; Gaps 4;

Qy 2312 GTGGCAGACGATGACCCAGGAGATATCAGTCTTGTCTGCTAGTCTTGTCTGAGCAGAGTGGAC 2371

Db 1 GGGGAGAGACGATGACCCAGGAGATATCAGTCTTGTCTGCTAGTCTTGTCTGAGCAGAGTGGAC 60

Qy 2372 AAGAAGCTTCCAGAACCTTGTCTTGGAGAAGTGTGATGAAGAAGATGAAGAAGTGTGACTTT 2431

Db 61 AAGAAGCTTCCAGAACCTTGTCTTGGAGAAGTGTGATGAAGAAGATGAAGAAGTGTGACTTT 120

Qy 2432 GGGAGGAAACAGGAGATTTGCTTACCTGAAGTGTGAGCCAAATCCAGAGAGACCCAGCAGTTT 2491

Db 121 GGGAGGAAACAGGAGATTTGCTTACCTGAAGTGTGAGCCAAATCCAGAGAGACCCAGCAGTTT 180

Qy 2492 ATCACCCTTCTACAGAGACTCTTGGCCCTTTGCTGGAGCCCTACAGCTCTGCTGCATC 2551

Db 181 ATCACCCTTCTACAGAGACTCTTGGCCCTTTGCTGGAGCCCTACAGCTCTGCTGCATC 240

Qy 2552 TTGTGTTCACAACTTTCAGTGTCTCTGTTTCCAGAACCTTGATATCTGCAAAAGTTGCAAAA 2611



```
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="Male and Female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/note="Vector: lambdaTriPEx"
BASE COUNT    173 a   145 c   131 g   184 t      4 others
ORIGIN
Query Match      18.8%; Score 564.2; DB 14; Length 637;
Best Local Similarity 98.4%; Pred. No. 4.7e-149;
Matches 599; Conservative 0; Mismatches 7; Indels 3; Gaps 3;
QY 1002 ATTTGGCGCTCTGTTCTACCGTTCATAGATCCCATATGCTATCTGCTCACTT 1061
Db 1 ATTTGGCGCTCTGTTCTACCGTTCATAGATCCCATATGCTATCTGCTCACTT 60
QY 1062 TCATTCTCTCTGCGATACCAATCAAGCAGCATATGCTTCAGGCAATAATCTCAACA 1121
Db 61 TCATTCTCTCTGCGATACCAATCAAGCAGCATATGCTTCAGGCAATAATCTCAACA 120
QY 1122 TCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTCATACGACGAAGGCTCG 1181
Db 121 TCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTCATACGACGAAGGCTCG 180
QY 1182 ATGAACACAGATGGAGGAGAGATGTTCTCTATAGAGCTTTCCTCCATGGGCATATAG 1241
Db 181 ATGAACACAGATGGAGGAGAGATGTTCTCTATAGAGCTTTCCTCCATGGGCATATAG 240
QY 1242 TTGAATTTACTTCGACAGCAGCAATTTCTGGAGATCTTCTCGAAGGACACAGTCTTAGGA 1301
Db 241 TTGAATTTACTTCGACAGCAGCAATTTCTGGAGATCTTCTCGAAGGACACAGTCTTAGGA 300
QY 1302 GTGGAATAAACCCTTTGTGCTCGGACGAGCTTTTGTGTCAGTGTGGTAGACTCTGCTA 1361
Db 301 GTGGAATAAACCCTTTGTGCTCGGACGAGCTTTTGTGTCAGTGTGGTAGACTCTGCTA 360
QY 1362 CCAATGTCATCCAGACATCTTGATATACCTGTTGGAATCTCTATGATCGCATATCG 1421
Db 361 CCAATGTCATCCAGACATCTTGATATACCTGTTGGAATCTCTATGATCGCATATCG 420
QY 1422 AAGTCTACTCAATGGTGAACAACTGGGCAACCTAAGAAGATGAGAGCTGTGGAGTG 1481
Db 421 AAGTCTACTCAATGGTGAACAACTGGGCAACCTAAGAAGATGAGAGCTGTGGAGTG 480
QY 1482 TAGCAAGAGGTGTATT-AGAATGTTACGAAAAAATATGTTGTGTGTCGAGTGGATTTT 1540
Db 481 TANCAGAGGTGTATTAAAGATGTTACGAAAAAATATGTTGTGTGTCGAGTGGATTTT 540
QY 1541 GCACAGCATTTTCTTAAAGGATATTTAAGACCAATGAGCAAGTCAAGACCGGTGCTGCT 1600
Db 541 GCACAGCATTTTCTTAAAGGATATTTAAGACCAATGAGCAAGTCAAGACCGGTGCTGCT 598
QY 1601 CTACTTTCC 1609
Db 599 CTACTTTCC 607

RESULT 11
BU703927
LOCUS
DEFINITION BU703927 753 bp mRNA linear EST 09-OCT-2002
IMAGE: 6406167 5', mRNA sequence.
ACCESSION BU703927
VERSION BU703927.1 GI:23631449
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 753)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```

## JOURNAL COMMENT

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgaabs@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA library preparation: Dr. M. Bento Soares, University of Iowa

cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. .753

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6"

/db\_xref="taxon:10090"

/clone="IMAGE: 6406167"

/tissue\_type="whole brain"

/dev\_stage="embryo 12.5dpc"

/lab\_host="DH10B (T1 phage resistant)"

/clone\_lib="NIH BMAP POO"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according

Bonaudo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is TAGAGAGCC. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator."

BASE COUNT 196 a 202 c 173 g 180 t 2 others

ORIGIN

Query Match 18.7%; Score 561.4; DB 13; Length 753;

Best Local Similarity 84.9%; Pred. No. 3.2e-148;

Matches 639; Conservative 0; Mismatches 113; Indels 1; Gaps 1;

QY 1373 CCAGACATCTTGATATACCTGTTGGAATCTCTATGATCGCATATCGAAGTCACTAC 1432

Db 1 CCGACATCTCGTCATACCCGCTGGGCATCTCGTATGATCGCATATCGAAGTCACTAC 60

QY 1433 AATGGTGAACAA-CTGGGCAACCTTAAGAAGATGAGAGCTGTGAGGTAGCAAGAGG 1491

Db 61 AATGGGCAACAAAGTTGGGAAAGCCCAAGAAGAACCGAGAGCTCTGGAGTGTGGCGAGAGG 120

QY 1492 TGTATTAGAAATGTTTACGAAAAAACTATGTTGTGTCGAGTGGATTTTTCACAGCCATT 1551

Db 121 CGTTATCAGATGCTCGGAAAAAACTACCGCTACGTCGAGTGGATTTTTCACAGCCATT 180

QY 1552 TTCTTTAAAGGAATATTTAGAAAGCCAAAGTCAAGAACCGGTGTCTGCTCTACTTTCCCT 1611

Db 181 TTCTTTAAAGGAATATTTAGAAAGCCCAAGTCAAGAACCGGTGTCTGCTCTACTTTCTCT 240

QY 1612 GGAGCAAGGTTGTTACGAGCTATCTTCTTCAAGACCCAGTGTGCTGCTGATCAAGG 1671

Db 241 GGAGCAAGCACTGTTTACCGCATCTCTCTTCAAGACCCGAATGATGTTGCTGATGAACA 300

QY 1672 TAGAGACAGCTCCATTATTAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGGAGGTT 1731

Db 301 TCAGACCTTATCCAGTACAGTCCAGAAACCCAGACAGCAAGCCCTTCGACGAAGGCT 360

QY 1732 GATTGCAATCTGGCTGAGCATATTTCTATTCTACTGCTAGCAAGTCTGTGCCATTATGTC 1791

Db 361 GATTGCAACCTGGCTGAGCATATTTCTTCTTCAACCGCAAGTCTGTGCCATTATGTC 420



QY 1792 CACACACATTGTGGCTGCTCTCTCTTACAGACAGGAGGAAATGATCTCTCCAC 1851  
 |||||  
 Db 421 CACCCACACATTGTGGCTGCTCTCTCTTACAGACAGGAGGAAATGATCTCTCCAC 480  
 |||||  
 QY 1852 ATTGGTGAAGACTCTTTTGTGATGAAGAGAACTCTGGCTCGTGATTTGACTGGG 1911  
 |||||  
 Db 481 GCTTGTGGAAAGACTCTTTTGTGATGAAGAGAACTCTGGCTCGTGATTTGACTGGG 540  
 |||||  
 QY 1912 GTTCTCAGGAAATTCAGAGATGTAGTAATGATGCCATACAGCTGCTGGGAAATTTGT 1971  
 |||||  
 Db 541 CTCTCCGGGAATTCAGAGATGTGCTCATGATGCTATTGAGCTTCTGGGGAATCTGT 600  
 |||||  
 QY 1972 CACAATCACCCACACTAGCAGGAACGATGAGTTTTTATFACCCCCAGCACAACCTGTCCC 2031  
 |||||  
 Db 601 CACAATCACCCACAGCAGGAGGAAAGATGAGTTTTTATFACCCCCAGCACAACCTGTCCC 660  
 |||||  
 QY 2032 ATCAGCTCTCGAATCACTTCTACGCAATGGGCTACTTCAATGCTTTATCATGAGGC 2091  
 |||||  
 Db 661 GTGATCTCTGAACTCAACTCTCAGCAATGGGCTACTTCAATGCTTTATCATGAGGC 720  
 |||||  
 QY 2092 CATCATAGCTTGCAGCTTTATGCACTTCTGAA 2124  
 |||||  
 Db 721 CATCATAGCTTGCAGCACTATGCACTCTGAA 753  
 |||||

RESULT 12  
 LOCUS BG704442 799 bp mRNA linear EST 07-MAY-2001  
 DEFINITION 602688632F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:4821342 5',  
 mRNA sequence.  
 ACCESSION BG704442  
 VERSION BG704442.1 GI:13977788  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC <http://mgi.nci.nih.gov/>.  
 1 (bases 1 to 799)  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM10727 row: o column: 07  
 High quality sequence stop: 704.

## FEATURES

source  
 1..799  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4821342"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 95"  
 /note="Organ: brain; Vector: pBluescriptR (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag  
 ); Oligo-dr primed using primer 5'-TTTTTTTTTTTNN-3',  
 size-selected for average insert size 2.5 kb and  
 normalized to ROT 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."  
 228 a 154 c 206 g 207 t 4 others

BASE COUNT

## ORIGIN

Query Match 18.5%; Score 555.6; DB 10; Length 799;  
 Best Local Similarity 93.4%; Pred. No. 1.5e-146;  
 Matches 654; Conservative 0; Mismatches 38; Indels 8; Gaps 7;  
 QY 213 GGCAGCAGCGGCTCCCTGTTGTATGGACATCTTCACCCAGCAGCATGTTGGGAATTACAC 332  
 |||||  
 Db 51 GGCAGCAGCGGCTCCCTGTTGTATGGACATCTTCGACCCGAACTGATAGTGAAGTCT 110  
 |||||  
 QY 273 GAAGCTTTATGTTATGAACAGAGAACTTTTCATCCAGCAGCATGATGTTGGGAATTACAC 332  
 |||||  
 Db 111 GAAGCTTTATGTTATGAACAGAGAACTTTTCATCCAGCAGCATGATGTTGGGAATTACAC 170  
 |||||  
 QY 333 TTTGTGACATGGATGAATCTGCACTGACCCCTTGGTACATAGATGTTCTTATCTGCCAC 392  
 |||||  
 Db 171 TTTGTGACATGGATGAATATGCACTGACCCCTTGGTACATAGATGTTCTTATCTGCCAC 230  
 |||||  
 QY 393 ATTCAATCAGATACAGTGTGGTGCATGTAAGCACACAAGTGAGGAAATGGGGTGAAGTGTG 452  
 |||||  
 Db 231 ATTCATCAGAATACAGTGTGGTGCATGTAAGCACACAAGTGAGGAAATGGGGTGAAGTGTG 290  
 |||||  
 QY 453 GCTTTAGACCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCTTAATGAGTC 512  
 |||||  
 Db 291 GCTTTAGACCCACCATCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCTTAATGAGTC 350  
 |||||  
 QY 513 GGAAGAGCCATTGTTGGAGATGTTGTTACTCTGCACCTCCAGAGCTGGGACAAAT 572  
 |||||  
 Db 351 GGAAGAGCCATTG-TGGAAGATGTTGTTACTCTGCACCTCCAGAGCTGGGACAAAT 409  
 |||||  
 QY 573 TTTTCAACCCCGATATCCCGTCTTTGGGTTTGGCGAATGTTTATTTATATCAATGAAATC 632  
 |||||  
 Db 410 TTTTCAACCCCGATATCCCGTCTTTGGGTTTGGCGAATGTTTATTTATATCAATGAAATC 469  
 |||||  
 QY 633 ACACAAGACACCGCGATGGCTTGCAGACGCCCTTTCTTACGTTCTTTTATTTCAAGAGC 692  
 |||||  
 Db 470 ACACAAGACACCGCGATGGCTTGCAGACGCCCTTTCTTACGTTCTTTTATTTCAAGAGC 529  
 |||||  
 QY 693 GAGATGTGCATAAGGGCATGTTTGCACCAATGACTGAAATGTCTGAACAGCAGTA 752  
 |||||  
 Db 530 GAGATGTGCATAAGGGCATGTTTGCACCAATGACTGAAATGTCTGAACAGCAGTA 589  
 |||||  
 QY 753 GAGTACAAGAGGCAATTCGAGAAGTGGCTGCTGAATTAACCCCTGATGTTCTGCCAGC 812  
 |||||  
 Db 590 GAGTACAAGAGG-AATTGCAGAGTGGTGTGTAATATAA-CCTGATGTTCTGCC--AG 645  
 |||||  
 QY 813 AGCAATCAAAAGCCGTTTAAACAAGTGAAGAAAGCTTAAAGGATTTCTTCAAGAAATGG 872  
 |||||  
 Db 646 CAGAATCAAAAGCCGTTTAAAC-AGTGAAGAAAGAGAG-TAAAGGATTTCTTC-AGAAATGG 702  
 |||||  
 QY 873 TTGCCACTGTCTCACCCGGCAATGATCAGACTGACTGGGTG 912  
 |||||  
 Db 703 TGGCACTGTTCTACGGGATGATCAGATGGATGGGTG 742  
 |||||

## RESULT 13

LOCUS AW976326 580 bp mRNA linear EST 02-JUN-2000  
 DEFINITION EST388435 MAGE resequences, MAGN Homo sapiens cDNA, mRNA sequence.  
 ACCESSION AW976326  
 VERSION AW976326.1 GI:8167552  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 580)  
 AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt  
 ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and  
 Quackenbush,J.  
 Assessment of gene expression patterns in a model of colon tumor  
 metastasis using a 19,200 element cDNA microarray  
 Unpublished  
 JOURNAL

Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: [johnq@tigr.org](mailto:johnq@tigr.org)  
Plate: 359

Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA

Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johnq@tigr.org  
Plate: 359

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	...	...	...	...
2	...	...	...	...
3	...	...	...	...
4	...	...	...	...
5	...	...	...	...
6	...	...	...	...
7	...	...	...	...
8	...	...	...	...
9	...	...	...	...
10	...	...	...	...
11	...	...	...	...
12	...	...	...	...
13	...	...	...	...
14	...	...	...	...
15	...	...	...	...
16	...	...	...	...
17	...	...	...	...
18	...	...	...	...
19	...	...	...	...
20	...	...	...	...
21	...	...	...	...
22	...	...	...	...
23	...	...	...	...
24	...	...	...	...
25	...	...	...	...
26	...	...	...	...
27	...	...	...	...
28	...	...	...	...
29	...	...	...	...
30	...	...	...	...
31	...	...	...	...
32	...	...	...	...
33	...	...	...	...
34	...	...	...	...
35	...	...	...	...
36	...	...	...	...
37	...	...	...	...
38	...	...	...	...
39	...	...	...	...
40	...	...	...	...
41	...	...	...	...
42	...	...	...	...
43	...	...	...	...
44	...	...	...	...
45	...	...	...	...
46	...	...	...	...
47	...	...	...	...
48	...	...	...	...
49	...	...	...	...
50	...	...	...	...
51	...	...	...	...
52	...	...	...	...
53	...	...	...	...
54	...	...	...	...
55	...	...	...	...
56	...	...	...	...
57	...	...	...	...
58	...	...	...	...
59	...	...	...	...
60	...	...	...	...
61	...	...	...	...
62	...	...	...	...
63	...	...	...	...
64	...	...	...	...
65	...	...	...	...
66	...	...	...	...
67	...	...	...	...
68	...	...	...	...
69	...	...	...	...
70	...	...	...	...
71	...	...	...	...
72	...	...	...	...
73	...	...	...	...
74	...	...	...	...
75	...	...	...	...
76	...	...	...	...
77	...	...	...	...
78	...	...	...	...
79	...	...	...	...
80	...	...	...	...
81	...	...	...	...
82	...	...	...	...
83	...	...	...	...
84	...	...	...	...
85	...	...	...	...
86	...	...	...	...
87	...	...	...	...
88	...	...	...	...
89	...	...	...	...
90	...	...	...	...
91	...	...	...	...
92	...	...	...	...
93	...	...	...	...
94	...	...	...	...
95	...	...	...	...
96	...	...	...	...
97	...	...	...	...
98	...	...	...	...
99	...	...	...	...
100	...	...	...	...

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi  
Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.  
1 (bases 1 to 602)  
Anson, W., Winkler, U., Mewes, H.W., Weil, B. and Wiemann, S.  
EST (Unpublished)  
Contact: Anson, W  
MIPS

Eukaryota; Metazoa; Chordata; Crania  
Mammalia; Euthera; Primates; Catarrhini  
1 (bases 1 to 602)  
Ansonge, W., Wirkner, U., Mewes, W., Weis-  
EST (Ansonge, W., Wirkner, U., Mewes, H.,  
Unpublished  
Contact: Ansonge W

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 602)  
Ansonge,W., Wirkner,U., Mewes,W., Weill,B. and Wiemann,S.)  
EST (Ansonge,W., Wirkner,U., Mewes,H.W., Weill,B. and Wiemann,S.)  
Unpublished  
Contact: Ansonge W

RESULT 15  
BG034520  
LOCUS  
DEFINITION BG034520 988 bp mRNA linear EST 24-JAN-2001  
602302638F1 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:4404225 5',  
mRNA sequence.  
ACCESSION BG034520  
VERSION BG034520.1 GI:12427920  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 988)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10115 row: k column: 10  
High quality sequence stop: 604.  
Location/Qualifiers  
1..988

FEATURES  
source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4404225"  
/tissue\_type="mammary adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 87"  
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.383 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH MGC Library."  
BASE COUNT 269 a 250 c 248 g 221 t

Query Match 17.6%; Score 530; DB 10; Length 988;  
Best Local Similarity 93.8%; Pred. No. 3.3e-139;  
Matches 619; Conservative 0; Mismatches 30; Indels 11; Gaps 6;  
Qy 1970 GTCACAATCACCACACTAGCAGGAACGATGAGTTTATACCCCGCAGCACAACTGTC 2029  
Db 1 GTCACAATCACCACACTAGCAGGAATGATGAGTTTATACCCCGCAGCACAACTGTC 60  
Qy 2030 CCATCAGTCTTGAACACTTCTACAGCAATGGGGTACTTCTTATCATGGAG 2089  
Db 61 CCATCAGTCTTGAACACTTCTACAGCAATGGGGTACTTCTTATCATGGAG 120  
Qy 2090 GCCATCATAGCTTGCAGCTTTATGAGTCTGTAACAGAGGGGACTGGGGGTCCCACT 2149  
Db 121 GCCATCATAGTCTGCAGCTTTATGAGTCTGTAACAGAGGGGACTGGGGGTCCCACT 180  
Qy 2150 AGCACCCACCTAACTGATCAGCCAGGAGCAGCTGGTGGCAAGGGCGCCAGCTGTGC 2209  
Db 181 AGCACCCACCTAACTGATCAGCCAGGAGCAGCTGGTGGCAAGGGCGCCAGCTGTGC 240  
Qy 2210 TACCTTCTCTCAATGAAGGACCACAT-CTCACTGCTTTGCCAGACATTTTACCAGTCTG 2268  
Db 241 TACCTTCTCTCAATGAAGGACCACATCTCACTGCTTTGCCAGACATTTTACCAGTCTG 300  
Qy 2269 CCATGAACAGTAGGAAGATTATCCAGTATGGGATTTCTTACAGTGGCAGACGATGA 2328  
Db 301 CCATGAACAGTAGGAAGATTATCCAGTATGGGATTTCTTACAGTGGCAGACGATGA 360

Qy 2329 CCAGGAAGATATCAGTCTCTAGTCTTGTGAGCAGTGGGCAAGAAGCTTCCAGAAC 2388  
Db 361 CCAGGAAGATATCAGTCTCTAGTCTTGTGAGCAGTGGGCAAGAAGCTTCCAGAAC 420  
Qy 2389 TTTGTCTTTGGAGAAGTGAAGAAGATGAAGACAGTGACTTTGGGGAGGAACAGCGAGA 2448  
Db 421 TTTGTCTTTGGAGAAGTGAAGAAGATGAAGACAGTGACTTTGGGGAGGAACAGCGAGA 480  
Qy 2449 TTGCTACTCTGAAGTGAAGCCAAAT-CCAAGGAGCACCAG-CAGTTTATCACCTTCTTACAG 2506  
Db 481 TTGCTACTCTGAAGTGAAGCCAAATCCAAGGAGCACCAGCCAGTTTATCAACCTTTCTTAC 540  
Qy 2507 AGA----CTCCTCTGGGCTTTGTGAGGCGCTACA--GCTCTGTGCGCATCTTTGTTTAC 2560  
Db 541 AGAGGACTCTCTTTGGGCTTTGTGAGGCGCTACAAGCTCTTGTGCGCATCTTTGTTTAC 600  
Qy 2561 AACTT--CAGTGTCTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTAA 2618  
Db 601 AACTTTCAGTGGTTCTCTGTTCCAGAACCTTGAATCTGCAAAAGTTGCAAAATACCTTA 660

Search completed: January 10, 2004, 06:10:19  
Job time : 4042.2 secs

**THIS PAGE BLANK (UNCLASSIFIED)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2004, 16:35:17 ; Search time 25 Seconds  
(without alignments)  
1557.524 Million cell updates/sec

Title: US-09-935-290-2

Perfect score: 4280

Sequence: 1 MDESALTGLTGDVSLPHSS.....FLPQCNKQLLEYILSFVVL 828

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4270	99.8	828	1	PLSB_HUMAN
2	3985.5	93.1	827	1	PLSB_MOUSE
3	3955	92.4	828	1	PLSB_RAT
4	678	15.8	718	1	PLSB_CABEL
5	505	11.8	870	1	PLSB_XYLEFA
6	462	10.8	834	1	PLSB_PSEAE
7	452	10.6	886	1	PLSB_XANCP
8	449	10.5	885	1	PLSB_XANAC
9	433	10.1	678	1	DAPT_RAT
10	424	9.9	678	1	DAPT_MOUSE
11	399.5	9.3	809	1	PLSB_VIBVU
12	398.5	9.3	808	1	PLSB_VIBPA
13	397	9.3	806	1	PLSB_SALTI
14	397	9.3	806	1	PLSB_SALTY
15	396	9.3	811	1	PLSB_VIBCH
16	395	9.2	810	1	PLSB_HAEIN
17	394	9.2	680	1	DAPT_HUMAN
18	391.5	9.1	806	1	PLSB_ECOLI
19	391	9.1	825	1	PLSB_ERFPE
20	390.5	9.1	806	1	PLSB_ECOS7
21	377	8.8	809	1	PLSB_PASMU
22	289	6.8	775	1	PLSB_MYCLE
23	245	5.7	789	1	PLSB_MYCTU
24	229	5.4	621	1	PLS1_MYCTU
25	140	3.3	3144	1	HD_HUMAN
26	128.5	3.0	967	1	CSE1_SCHPO
27	124.5	2.9	3110	1	HD_RAT
28	123.5	2.9	1010	1	UF22_SCHPO
29	119.5	2.8	1174	1	YJ11_YEAST
30	118.5	2.8	1176	1	RFB2_DROME
31	118	2.8	1258	1	SAL1_HUMAN
32	118	2.8	1258	1	SAL1_MOUSE
33	116.5	2.7	958	1	BP28_MACFA

#### ALIGNMENTS

##### RESULT 1

ID	PLSB_HUMAN	STANDARD;	PRT;	828 AA.
AC	Q9HCL2;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glycerol-3-phosphate acyltransferase, mitochondrial precursor			
DE	(EC 2.3.1.15) (GPAT).			
GN	GPAT OR KIAA1560.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Ashwell R.;			
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE OF 167-828 FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=20450683; PubMed=10997877;			
RA	Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes.			
RT	XVIII. The complete sequences of 100 new cDNA clones from brain which			
RL	code for large proteins in vitro.";			
RL	DNA Res. 7:273-281(2000).			
CC	-!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-			
CC	acyl-en-glycerol 3-phosphate.			
CC	-!- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS. IT MAY			
CC	ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial			
CC	outer membrane (By similarity).			
CC	-!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; AL391986; -; NOT ANNOTATED_CDS.			
DR	EMBL; AB046780; BAB13356.1; -;			
DR	MIM; 602395; -;			
DR	InterPro; IPR001213; Acyltransferase.			
DR	Pfam; PF01553; Acyltransferase; 1.			
DR	SMART; SM00563; Plac; 1.			
KW	Phospholipid biosynthesis; Transferase; Acyltransferase;			
KW	Transmembrane; Mitochondrion; Transit peptide.			
FT	TRANSIT 1 ?			
FT	CHAIN ?			
FT	DOMAIN ?			
FT	TRANSMEM 472 494			
FT	DOMAIN 495 574			

34	115.5	2.7	3119	1	HD_MOUSE	P42859	mus musculus
35	115	2.7	3744	1	YHP9_YEAST	P38811	saccharomyc
36	111	2.6	987	1	SA2_YEAST	P16861	saccharomyc
37	111	2.6	1162	1	SA2_HUMAN	O8n3u4	homo sapien
38	110.5	2.6	820	1	MUTS_CHLTR	O84797	chlamydia t
39	110.5	2.6	1024	1	GCPS_HUMAN	Q96rt8	homo sapien
40	109	2.5	1235	1	KFB2_RABIT	P46018	oryctolagus
41	108.5	2.5	768	1	LIPS_RAT	P15304	rattus norv
42	108	2.5	520	1	GUAA_RHILO	Q987r3	rhizobium l
43	108	2.5	814	1	AKH_BUCAP	O8k9u9	buchnera ap
44	108	2.5	1162	1	SA2_MOUSE	O35638	mus musculus
45	108	2.5	2273	1	HPA1_YEAST	P32874	saccharomyc

```

FT TRANSMEM 575 593 POTENTIAL.
FT DOMAIN 594 828 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
SQ SEQUENCE 828 AA; 93836 MW; 55CD321A23D0B65B CRC64;

Query Match
Best Local Similarity 99.8%; Score 4270; DB 1; Length 828;
Matches 826; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDSALTLGTIDVSYLPHSSEYSGVGRCKHTSEWEGCGFRPTVFRSATLKWESLSMRKR 60
DB 1 MDSALTLGTIDVSYLPHSSEYSGVGRCKHTSEWEGCGFRPTVFRSATLKWESLSMRKR 60
QY 61 PFVGRCCYCTPQSDWKDFNPSISGLRNVIYINETHRHRGLARLSYVLFQERDV 120
DB 61 PFVGRCCYCTPQSDWKDFNPSISGLRNVIYINETHRHRGLARLSYVLFQERDV 120
QY 121 HKGMFATNTENVLNSSRVQEAIAEVAELNPDGSAQOQSKAVNKKAKRILQEMVAT 180
DB 121 HKGMFATNTENVLNSSRVQEAIAEVAELNPDGSAQOQSKAVNKKAKRILQEMVAT 180
QY 181 VSPAMIRLTGWLLKLFNSFFWNIQHKGLEWVKAATETNPLFLPVHRSHIDYLLT 240
DB 181 VSPAMIRLTGWLLKLFNSFFWNIQHKGLEWVKAATETNPLFLPVHRSHIDYLLT 240
QY 241 FILFCHNIKAPYIASGNLNIPFSTLIHLKGGFFIRRRDETPDGRKDVLYRALLHGH 300
DB 241 FILFCHNIKAPYIASGNLNIPFSTLIHLKGGFFIRRRDETPDGRKDVLYRALLHGH 300
QY 301 VELLRQOQFLEIFLEGTRSRGKTSACAGLLSVVDLTSTNVIPIIDLIIPVGISYDR 360
DB 301 VELLRQOQFLEIFLEGTRSRGKTSACAGLLSVVDLTSTNVIPIIDLIIPVGISYDR 360
QY 361 ECHYNGEQLKPKKNESLWSVARGVIRMLRKNVGCVRVDFAPFSLKEYLESQKPVSA 420
DB 361 ECHYNGEQLKPKKNESLWSVARGVIRMLRKNVGCVRVDFAPFSLKEYLESQKPVSA 420
QY 421 LLSLEQALLPAILPSRSDAADERGTSINESRNATDESRLRLIANLAELHILFTASKSC 480
DB 421 LLSLEQALLPAILPSRSDAADERGTSINESRNATDESRLRLIANLAELHILFTASKSC 480
QY 481 AIMSTHIVACLLYRHRQGDLSLTVEDFFVMKEEVLARDFDLFGSGNSEDDVVMHAIQL 540
DB 481 AIMSTHIVACLLYRHRQGDLSLTVEDFFVMKEEVLARDFDLFGSGNSEDDVVMHAIQL 540
QY 541 GNCVTITHTSRNDEFFITPTSPVPELNFYNGVLHVIMEAIIACSLYAVLNKRGGLG 600
DB 541 GNCVTITHTSRNDEFFITPTSPVPELNFYNGVLHVIMEAIIACSLYAVLNKRGGLG 600
QY 601 GPTSTPNLLISQBLVRKAASLCYLLSNEGTSISLPCQTFYQVCHETVKGFIQIGILTVAE 660
DB 601 GPTSTPNLLISQBLVRKAASLCYLLSNEGTSISLPCQTFYQVCHETVKGFIQIGILTVAE 660
QY 661 HDQOEDISPLAEOQMDKKLPEPLSWRSDEDESDSDFGEEQRDCYLKVSQKEHQOIF 720
DB 661 HDQOEDISPLAEOQMDKKLPEPLSWRSDEDESDSDFGEEQRDCYLKVSQKEHQOIF 720
QY 721 LQRLGLPLLAYSAALFVNFNPGSPVPEPEYLOKLHKYILTRERNVAVVAESATYCLVK 780
DB 721 LQRLGLPLLAYSAALFVNFNPGSPVPEPEYLOKLHKYILTRERNVAVVAESATYCLVK 780
QY 781 NAVMFKDIGVFKETKOKRVSVLELSSTFLPQCNQKLLVILSVVL 828
DB 781 NAVMFKDIGVFKETKOKRVSVLELSSTFLPQCNQKLLVILSVVL 828

```

## RESULT 2

```

ID_PLSB_MOUSE STANDARD; PRT; 827 AA.
AC Q61586;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycerol-3-phosphate acyltransferase, mitochondrial precursor

```

```

DE (BC 2.3.1.15) (GPAT) (P90).
GN GPAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92084678; PubMed=1721057;
RA Shin D.-H., Paulauskis J.D., Moustaid N., Sul H.S.;
RT "Transcriptional regulation of p90 with sequence homology to
RT Escherichia coli glycerol-3-phosphate acyltransferase.";
RL J. Biol. Chem. 266:23834-23839(1991).
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS. IT MAY
CC ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC outer membrane (By similarity).
CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN LIVER, INTERMEDIATE LEVELS
CC IN MUSCLE AND KIDNEY, AND LOWEST LEVELS IN LUNG AND BRAIN.
CC -1- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M77003; AAA37647.1; -.
DR PIR; A41672; A41672.
DR MGD; MGI:109162; Gpat.
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
DR SMART; SM00563; Plac; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Transmembrane; Mitochondrion; Transit peptide.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN 827 GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
FT DOMAIN ? 471 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT TRANSMEM 472 494 POTENTIAL.
FT DOMAIN 495 574 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 575 593 POTENTIAL.
FT DOMAIN 594 827 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
SQ SEQUENCE 827 AA; 93689 MW; A64EDE697BD664B3 CRC64;

Query Match 93.1%; Score 3985.5; DB 1; Length 827;
Best Local Similarity 92.5%; Pred. No. 1.2e-284;
Matches 766; Conservative 32; Mismatches 29; Indels 1; Gaps 1;

QY 1 MDSALTLGTIDVSYLPHSSEYSGVGRCKHTSEWEGCGFRPTVFRSATLKWESLSMRKR 60
DB 1 MEBSVTVGTIDVSYLPHSSEYSLGCKHTSEWVDCGKPTFRSATLKWESLSMRKR 60
QY 61 PFVGRCCYCTPQSDWKDFNPSISGLRNVIYINETHRHRGLARLSYVLFQERDV 120
DB 61 PFVGRCCYCTPQSDWKDFNPSISGLRNVIYINETHRHRGLARLSYVLFQERDV 120
QY 121 HKGMFATNTENVLNSSRVQEAIAEVAELNPDGSAQOQSKAVNKKAKRILQEMVAT 180
DB 121 HKGMFATNTENVLNSSRVQEAIAEVAELNPDGSAQOQSKAVNKKAKRILQEMVAT 180
QY 181 VSPAMIRLTGWLLKLFNSFFWNIQHKGLEWVKAATETNPLFLPVHRSHIDYLLT 240
DB 181 VSPAMIRLTGWLLKLFNSFFWNIQHKGLEWVKAATETNPLFLPVHRSHIDYLLT 240
QY 241 FILFCHNIKAPYIASGNLNIPFSTLIHLKGGFFIRRRDETPDGRKDVLYRALLHGH 300
DB 241 FILFCHNIKAPYIASGNLNIPVFTLIHLKGGFFIRRRDETPDGRKDVLYRALLHGH 300

```

```
QY 301 VELLRQOQFLEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPDILIPVGSYDRII 360
Db 301 VELLRQOQFLEIFLEGTRSRSGKTSARAGVLSVVVNTLSSNTIPDILIPVGSYDRII 360
QY 361 EGHYNGEOLGPKKNESLWSVARGVIRMLRKNYGVVDFAQPFSLKYLESSQSKPVSA 420
Db 361 EGHYNGEOLGPKKNESLWSVARGVIRMLRKNYGVVDFAQPFSLKYLESSQSKPVSA 420
QY 421 LLSLEQALLPAILPRSPDADEGRDTSINESRNATDESRLRRLLIANAEHLFTASKSC 480
Db 421 PLSLEQALLPAILPRSPNDVADEHQDLSNERNPADAEFRRLIANAEHLFTASKSC 480
QY 481 AIMSTHIVACLLYHRQIGIDLTSLVDFVFMKEEVLARDPDLGFGSGNSEDVWMAIQLL 540
Db 481 AIMSTHIVACLLYHRQIGIHLSTLVEDFVFMKEEVLARDPDLGFGSGNSEDVWMAIQLL 540
QY 541 GNCVTITHTSRNDEFFITPTVPSPFELNFSYNGVLHVFMEALIIACSLVAVLNKRGIG 600
Db 541 GNCVTITHTSRNDEFFITPTVPSPFELNFSYNGVLHVFMEALIIACSLVAVLNKRGIG 600
QY 601 GPTSTPPNLSQELVRKAASLCYLLSNEGTLISLPCQTFYQVCHETVGVKFIQYGLITVAE 660
Db 601 GSAGLGNLSQELVRKAASLCYLLSNEGTLISLPCQTFYQVCHETVGVKFIQYGLITVAE 660
QY 661 HDDQEDISPLAEOQWKKLPEPLSWRSDDEDESDPGEOQDCYLKVSQKEHQOQFITF 720
Db 661 QDDQEDVSPGLAEQWKKLPE-LNWRSDDEDESDPGEOQDCYLKVSQKEHQOQFITF 719
QY 721 LQRLGLPILLEYSSAAIFVHNFSGVPPEYLOKLHKYLIIRTERNAVVAESATYCLVK 780
Db 720 LQRLGLPILLEYSSAAIFVHNFSGVPPEYLOKLHKYLIIRTERNAVVAESATYCLVK 779
QY 781 NAVMKFDIGVFKETKQKRVSVLESLSTFLPCQNRQKLEVLILSPVWL 828
Db 780 NAVMKFDIGVFKETKQKRVSVLESLSTFLPCQNRQKLEVLILSPVWL 827
RESULT 3
PLSB RAT STANDARD PRT; 828 AA.
AC P97564; Q35349; P97565; P97566;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Glycero1-3-phosphate acyltransferase, mitochondrial precursor
DE (EC 2.3.1.15) (GPAT).
CN GPAM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND ACTIVITY.
RC TISSUE=Liver;
RX MEDLINE=99376617; PubMed=1046428;
RA Bhat B.G., Wang P., Kim J.-H., Black T.M., Lewin T.M.,
RA Fiedorek F.T. Jr., Coleman R.A.;
RT "Rat sn-glycerol 3-phosphate acyltransferase: molecular cloning and
RT characterization of the cDNA and expressed protein.";
RL Biochim. Biophys. Acta 1439:415-423(1999).
RN [2]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TOPOLOGY.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=20493538; PubMed=10924502;
RA Balija V.S., Chakraborty T.R., Nikonov A.V., Morimoto T., Haldar D.;
RT "Identification of two transmembrane regions and a cytosolic domain of
RT rat mitochondrial glycerophosphate acyltransferase.";
RL J. Biol. Chem. 275:31668-31673(2000).
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -!- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS. IT MAY
CC ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
```

```
outer membrane.
-!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF021348; AAB71605.1; -
DR EMBL; U36771; AAB39470.2; ALT INIT.
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
DR SMART; SM00563; Plsc; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Transmembrane; Mitochondrion; transit peptide.
FT TRANSIT 1 ?
FT CHAIN ? 828 MITOCHONDRION (POTENTIAL).
FT DOMAIN ? 471 GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
FT TRANSMEM 472 494 POTENTIAL.
FT DOMAIN 495 574 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 575 593 POTENTIAL.
FT DOMAIN 594 828 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT CONFLICT 37 37 C -> F (IN REF. 2).
FT CONFLICT 85 85 S -> P (IN REF. 2).
FT CONFLICT 300 300 I -> V (IN REF. 2).
FT CONFLICT 331 331 L -> V (IN REF. 2).
FT CONFLICT 472 474 ILF -> NLL (IN REF. 2).
FT CONFLICT 497 497 R -> W (IN REF. 2).
FT CONFLICT 602 606 SAGGL -> LPEP (IN REF. 2).
FT CONFLICT 644 644 Q -> H (IN REF. 2).
FT CONFLICT 744 744 G -> A (IN REF. 2).
SQ SEQUENCE 828 AA; 93714 MW; ACA4A087E7DEB12C CRC64;
Query Match 92.4%; Score 3955; DB 1; Length 828;
Best Local Similarity 91.7%; Pred. No. 2.1e-282;
Matches 759; Conservative 34; Mismatches 35; Indels 0; Gaps 0;
QY 1 MDESALTGTITDVSYLPHSSSEYSGRCCKHTSEEWGECGRPTVFRSATLKWESLMSKR 60
Db 1 MEESVTTGTTDVSYLPHNSSEYSLGRCKHTNEDWDCGKPTFRSATLKWESLMSKR 60
QY 61 PFVGRCCVCTPQSQWDXFFNPSIPSLGRNVIYINETHRHRGMLARLSYLVFTQERDV 120
Db 61 PFVGRCCVCTPQSQWERFFNPSIPSLGRNVIYINETHRHRGMLARLSYLVFTQERDV 120
QY 121 HKGMFATNVTENVLSNRVQEAIAEVAELNPDGSAQOQSKAVNKKKAKRILOEMVAT 180
Db 121 HKGMFATNVTENVLSNRVQEAIAEVAELNPDGSAQOQSKAIQKVKEKARKILOEMVAT 180
QY 181 VSPAMIRLTGWVLLKLFNSFPWNTQIHKGQLEWYKAATETNLLPLFVHRSHIDYLLLT 240
Db 181 VSPGMIRLTGWVLLKLFNSFPWNTQIHKGQLEWYKAATETNLLPLFVHRSHIDYLLLT 240
QY 241 FILFCHNIKAPYIAGNNLNPIESTLHKLGGSFIRRLDETDPGRKDVYRALLHGH 300
Db 241 FILFCHNIKAPYIAGNNLNPIESTLHKLGGSFIRRLDETDPGRKDVYRALLHGH 300
QY 301 VELLRQOQFLEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPDILIPVGSYDRII 360
Db 301 VELLRQOQFLEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPDILIPVGSYDRII 360
QY 361 EGHYNGEOLGPKKNESLWSVARGVIRMLRKNYGVVDFAQPFSLKYLESSQSKPVSA 420
Db 361 EGHYNGEOLGPKKNESLWSVARGVIRMLRKNYGVVDFAQPFSLKYLESSQSKPVSA 420
QY 421 LLSLEQALLPAILPRSPDADEGRDTSINESRNATDESRLRRLLIANAEHLFTASKSC 480
Db 421 PLSLEQALLPAILPRSPDADEGRDTSINESRNATDESRLRRLLIANAEHLFTASKSC 480
QY 481 AIMSTHIVACLLYHRQIGIDLTSLVDFVFMKEEVLARDPDLGFGSGNSEDVWMAIQLL 540
```



Db 481 AINSTHIVACLLYRHRQGIHLSTLVDFPMKEEVLARDFDLGFGNSDVMHAIQLL 540  
Qy 541 GNCVTHTRSRNDEFTPTPTTPSPVFNLYNGVHLVMEALITACSLYAVLNKRGIG 600  
Db 541 GNCVTHTRSRNDEFTPTPTTPSPVFNLYNGVHLVMEALITACSLYAVLNKRGIG 600  
Qy 601 GPTSTPNLISQOLVRKAASLCVLLSNEGTISLPQOTFYOVCHETVGRFIOVGLITVAE 660  
Db 601 GSAGGLNLISQOLVRKAASLCVLLSNEGTISLPQOTFYOVCHETVGRFIOVGLITVAE 660  
Qy 661 HDQEDISPLAQQMDKKLPEPLSRWSDEDESDFGEEQRDCYLKVSQSKHQQFITF 720  
Db 661 QDQEDVSPGLAQWKKLPEPLNRWSDEDESDFGEEQRDCYLKVSQSKHQQFITF 720  
Qy 721 LQRLGLLEAYSSAIFVNFSPGVPPEPYLOKLHLYLTRNVAIVAAATYCLVK 780  
Db 721 LQRLGLLEAYSSAIFVNFSPGVPPEPYLOKLHLYLTRNVAIVAAATYCLVK 780  
Qy 781 NAVKMFQDGVFKETKQKRVSVLELSSTFLPQCNROKLLVILSFVVL 828  
Db 781 NAVKMFQDGVFKETKQKRVSVLELSSTFLPQCNROKLLVILSFVVL 828

## RESULT 4

PLSB CAEEL STANDARD; PRT; 718 AA.  
AC Q22949;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Probable glycerol-3-phosphate acyltransferase, mitochondrial precursor  
DE (EC 2.3.1.15) (GPAT).  
GN F08F3.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Blanchard M., Bradshaw H.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
CC -I- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-  
CC acyl-sn-glycerol 3-phosphate.  
CC -I- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS. IT MAY  
CC ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.  
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
CC (Potential).  
CC -I- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U64847; AAB04876.1; -.  
DR PIR; T29448; T29448.  
DR WormPep; F08F3.2; CE09258.  
DR InterPro; IPR002123; Acyltransferase.  
DR Pfam; PF01553; Acyltransferase; 1.  
DR SMART; SM00563; Plac; 1.  
KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
KW Transmembrane; Mitochondrion; Transit peptide.  
FT TRANSIT 1 ?  
FT CHAIN ? 718 MITOCHONDRION (POTENTIAL).  
FT PROBABLE GLYCEROL-3-PHOSPHATE  
FT ACYLTRANSFERASE.  
FT  
FT TRANSMEM 409 425  
FT SEQUENCE 718 AA; 82071 MW; E0A36A4A86FC138D CRC64;  
SQ

Query Match 15.8%; Score 678; DB 1; Length 718;  
Best Local Similarity 27.2%; Pred. No. 6.8e-42;  
Matches 197; Conservative 148; Mismatches 264; Indels 114; Gaps 21;  
Qy 128 NTVNVLNRSRVOEAEVAELNPDGSAQQSKAVNKKAKRILQEWVATVSPAMIR 187  
Db 78 NVEKDVLSKRVRHVISKL-----KEQN---DEQNRVQVFTTISARLSFKICK 124  
Qy 188 LTGWLLKLFNSFFWNIQIHKGQLEVMKAAETNLPLLFVHRSHIDVLLTFLFCHN 247  
Db 125 CCSYVLVYKFRMLMDKLLVCKEEMEVYEAQGTIPMVYLLPHRSHLDVLLITCNWHFG 184  
Qy 248 IKAPYIASGNLNIPIFTLIHKLGGFIRRRIDETPDGRKDVLYRALLHGHIVELLRQ 307  
Db 185 LKLPHTASGDNLSGLWLLRATGATGAFRRVDPDERGKQLYRALLSHYIEQVLSKD 244  
Qy 308 QFLEIPLEGTRSRSGKTS CARAGLLSVVVDLTSTNVIPDILIPVGISVDRIIEGHNGE 367  
Db 245 MPTIEFFLEGTRSRFGKALTPKNGLISNVVVEAQHGFIKDCYLPVPSYTYDAVVEGIFLHE 304  
Qy 368 QLGKPKKNEFSLMSVARGVIRMLRNKNGC--VRVDFAPQPSLKEYLESQSKPVSALLSLE 425  
Db 305 LMGIPKVRSEVLGVRGIFSGFSKQCGVWRMHYGRPIRLTEYLAT-----ITASLSSN 359  
Qy 426 QALLPAILPSRPSDAADGRDTSINES-----RNATDESRRRLIANIAEHILEFT 475  
Db 360 HRTRPV-----RMTKLSFSYSRELVPMHRTHTSETVDDRTWIRALGFHVVE 406  
Qy 476 ASKSCAIMSTHIVACLLYRHRQGIHLSTLVDFPMKEEVLARDFDLGFGNSDVMHAIQLL 531  
Db 407 AQWMCSSIPVAVVSCILLAKWRGKVSRTFERDCWLCCKIIAEGGDVVGYQSKTKGSA 466  
Qy 532 VVMAHQLLGNCVTIHTSRNDEFTPTPTTPSPVFNLYNGVHLVMEALITACSLY 591  
Db 467 LVKYAFKLEKSEVEVT-----DE-VVSPKSHSSFTITAYNKNVTCRFSIKSVIALTI- 519  
Qy 592 AVLNRKGLGGTSTPPNLLISQOLVRKAASLCVLLSNEGTISLPQOTFYOVCHETVGRKFI 651  
Db 520 -----VSRPSTGK---LSIQIVEDALSCLDWLQFECPCPCDLSRELNVNLG--- 566  
Qy 652 QYGILTVAEHDDQEDISPLAQQMDKKLPEPLS--WRSEDEDED--SDGEEORDCYLVK 708  
Db 567 -----OKEWS-----DPIHGFRLSEIEDDGFGLDAGGALNSGTLRV 601  
Qy 709 SQSKEHQQFITFLQRLGLLEAYSSAIFVNFSPGVPPEPYLOKLHLYLTR----- 763  
Db 602 RDAKS--RETLPFANLVRPFVQSLYLISSFVSEKCP--EPTSDNNIIRQLCQOQSLAGDI 659  
Qy 764 ERNVAVYAESATYCLVKNVAKMFKDIGVFKETKQKRV-----VLELSSTFLPQCNROK 817  
Db 660 DLPFAPLESINSDSFKNALRVLKDKGLLQRTSPNSTARSGNSRLAELIS-----NLER 713  
Qy 818 LLE 820  
Db 714 VLE 716  
RESULT 5  
PLSB XYLFA STANDARD; PRT; 870 AA.  
AC Q9PEJ37;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).  
GN PLSB OR XF1031.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xylella.  
OX NCBI\_TaxID=23711;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9a5c;

RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britones M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Canagar L.E.A., Carraro D.M., Carrier H.,  
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neco C.M.,  
 RA Coutinho L.P., Cristofani M., Dias-Neco E., Docena C., El-Dorry H.,  
 RA Facincini A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
 RA Garnier J.M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hohlseil J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kuramse E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Mend C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.B. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
 RA "The genome sequence of the plant pathogen Xylella fastidiosa";  
 RT Nature 406:151-159(2000).  
 CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-  
 CC acyl-sn-glycerol 3-phosphate.  
 CC -!- PATHWAY: DE NOVO PHOSPHOLIPID BIOSYNTHESIS; FIRST STEP. MAY ALSO  
 CC FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AE003940; AAF83841.1; --  
 DR PIR; B82732; B82732.  
 DR HAMAP; MF\_00393; --; 1.  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 DR SMART; SM00563; PlsC; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;  
 KW Complete proteome.  
 SQ SEQUENCE 870 AA; 98483 MW; 52A8ADA99FB9868 CRC64;  
 Query Match 11.8%; Score 505; DB 1; Length 870;  
 Best Local Similarity 22.4%; Pred. No. 4.5e-29;  
 Matches 197; Conservative 148; Mismatches 314; Indels 222; Gaps 26;  
 QY 13 VSYLPHSEYSVGR-CKHTSEWGC-----GPRPTVFSATLWKESLSMRKRPVVG 64  
 DB 64 VOYPEHCVOYNDGPIIYVLEDYGLCNTLILDKACRKTLPSPILPFGNLPQKRAYLA 123  
 QY 65 RCYCYCT-----PQSWDKFFNPSISPL-----GLRNV-----IYNETHTRHGWLA 106  
 DB 124 LSRRSSNSLIPNQGGKTHSDSLANLQAHRIITDLVHLVPVSIFIGRTPDROSGWFA 183  
 QY 107 -----RRLSYVLPIQ 116  
 DB 184 VLFSENWALVGRFRELAVLLNGRNTIVCFAPPISVQRTLNEGLPPERTLKLQVLRH 243  
 QY 117 ERDVHKGFAFNTV-----ENVLSRVQEAIAEVAALPDGSAQQOQKAVNKKVKA 170  
 DB 244 FRRIRETVIGPDLSTRLLVDVNLATEAVREAIA-----SQAKRDGTDLSWTWKA 294

QY 171 KRILQEWVATVSPAMIRLTGWLLKLKLFNSFFWNIIQHKQLEMKAAATETNLPILFLPVH 230  
 DB 295 QAYAEIAADYSSPVIRSADFLFHVWNRIVAGVLIIH--HVDSEK-ETAPGHEVVYVESH 351  
 QY 231 RSHIDYLLLTFLFCHNIIKAPYIASGNINIIPISTLIIKLGPGFIRRLRDLTPDGRKDV 290  
 DB 352 RSHIDYLLSYCLQCQGVILPHIVAGINLNLPIVGTLLRKCCAFIRRSI-----KGNM 405  
 QY 291 LYRALLGHIVELARQQQFLIELEGTSRSGKTSACAGLLSVVVDTLSTNVIPDILII 350  
 DB 406 LYSIVLSEYVAQLVAGGYSLEYFIEGSRSTRGLLPQKGGIMMTLQAFLOKQPRRPVLFQ 465  
 QY 351 PVGISYDRIIEG-HYNGEQLGPKKNESLWSVARGVIRMLRKNYGCVRVDVDAQPSLKEY 409  
 DB 466 PIYIGYEKLIBGTSYLDSESGPKKESIMELFNWIPKVLAKYQGVVVNFGEPALANDV 525  
 QY 410 LESQSKPVSALESQALLPAILPSRPSDAADSGRDTSSINESRNATDESILRRLLIANLA 469  
 DB 526 L-----AELAPEWEGQALNENKPAWLS---NTVNHLLA 555  
 QY 470 EHILFTASKSCAIMSTHIVACILLYRHQGDIDLTSLVEDFFVMKEEVLARDDFDLGFSG-- 527  
 DB 556 ROIQTRINSAADVNPINLLALLSTPKHAMEADLITQAQITLCKKILL-----ELPYSNRV 611  
 QY 528 ----NSDDVVMHATQLIGNCVITITHSRNDDEFFITPTTVPSPFELNFYNGVLHVIFIM 582  
 DB 612 TVTHTPERIATAHAEQI--NLITRVHHPGLDVLVDGDNV-----LLSYFNNVHLHFTA 665  
 QY 593 EAITACSLYAVLNKRGGLGPTTPNLIISQOLVRKAASLCYLLSNETGISLPCQTFYQV 642  
 DB 666 SAWVACCFKN--NRR-----ISRILRLGVGYFPFLQAEFLPWTEDQFAQH 711  
 QY 643 CHETVGKFIQVIGILTVAEHDDQEDISPSLAEQWQDKLPELSWRSDEDESDPGEQR 702  
 DB 712 IQQVIELFVREGLL--LSAGDEED-----PLTRNTSQTDE----- 745  
 QY 703 DCYLKVSQSKHQHQQFIPTLQRLGLGLEA-----YSSAAIFVHNPFGVPPEYLQKLHK 757  
 DB 746 -----VFLRAISHSQOAFERYITISILVKNPGTILSASE-----LES 785  
 QY 758 YLITRTERNVAVYAESATYCLVKNVAKMFKDGVFKETKQK 798  
 DB 786 LCQLAAQRLSLLYASTA-----PEFFDKGLFRGFQIK 817  
 RESULT 6  
 ID PLSB\_PSEAE STANDARD; PRT; 834 AA.  
 AC Q9HXW7;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).  
 GN PLSB OR PA3673.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 15692 / PAOI;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RA "Complete genome sequence of Pseudomonas aeruginosa PAOI, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-

CC acyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: DE NOVO PHOSPHOLIPID BIOSYNTHESIS; FIRST STEP. MAY ALSO  
CC FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.  
CC -1- SUBCELLULAR LOCATION: Membrane-bound (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC  
CC EMBL: AE004787; AAC07061.1; --  
CC PIR: F83185, F83185.  
CC HAMAP: MF\_00393; -, 1.  
CC InterPro: IPR002123; Acyltransferase.  
CC Pfam: PF01553; Acyltransferase; 1.  
CC SMART: SM00563; Pfam; 1.  
CC Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;  
CC Complete proteome.  
CC SEQUENCE 834 AA; 94815 MW; C668FB7242B3681D CRC64;  
Query Match 10.8%; Score 462; DB 1; Length 834;  
Best Local Similarity 22.8%; Pred. No. 6.1e-26;  
Matches 171; Conservative 150; Mismatches 262; Indels 166; Gaps 27;  
QY 88 LRVVYINETHRRGWLRLSVLPFIQDRVHKMFA-----TNVTNLSNRVQ 141  
DB 178 LRELVEQKQHER-----TLRMVNRILVHFRNLKTAIVGPDLSHRRLNVLKGLRAPLVRQ 233  
QY 142 AIAEVAELNPDGSAQOQSAVNVKVKKAKILQEMVATVSPAMIRLTGWVLKLFNSFF 201  
DB 234 AISE-----ECSEIRISQKAE-----ALRYANETIASDFSYPVIR-----FLEVILSWF 279  
QY 202 WNTQIHKG-----QLEMKAAATETNPLPLFVHRSHIDYLLLFILFCHNIKAPYIASGN 257  
DB 280 WN-KLYEGVKVNIHIERQDVQAGN-EIVYVPCHRSHIDYLLSVLLFRNGLTPEPHIAAGI 337  
QY 258 NLNIPFSTLIHLKGGFFIRRLDETDPGRKDVLYRALLHGHIVELLRQQOFTLEIFLEGT 317  
DB 338 NLNMPVIGSLRRGGAFMRSE-----KGNOLYTAFFNEYLTLFSGFSTFYEVEGG 391  
QY 318 RSRGKTSRAGLSVVVDTLSNVIPDILLIPVGSYDRILEGH-YNGEQKGPKN 376  
DB 392 RSTGRMLHPTGMLATLRSFDRSRPVPVPIYIGERVLEGRTYLGEKATKK 451  
QY 377 SLWSVARGVIRMLRKNYGVVDFAPQPSLKEYLESQSPVSAALLSLEQALLPAILPS- 435  
DB 452 SIFDLFK-VVGALKQRFQGVVNFGEPIHLDQFLDRHQ-----PDW 491  
QY 436 RPSDADEGRDTSINESRNATDESRLRLIANLAHLFTASKSCAMSTHIVACLLLYR 495  
DB 492 QDQDLGPEYRPMPLPQTNL-----LAKDVARHL-----NDAAALNPVNLVALALLST 539  
QY 496 HRQIDLSLTVDFEYFVKMEVILARDFDLGS-----GNSDEVVHAIQLL 540  
DB 540 SRQALD-----ESALRILDLYLALRKVPYSPSATLPDGGQALLYEVKSM- 586  
QY 541 GNCVTITHTSRNDEFFITPSTVPVSFELNFYSNGVLHVFIMBAITACSLYAVLNKRG 600  
DB 587 -LLAEQKDALGRILYLDQNAVLA-----TYRRNVLVHVPALPALIASFFQS-NSR- 636  
QY 601 GPTSTPPNLSIQQLVKAASLCVLLSNEGTSILPCTQFTVQVCHETVKGFIQVILTVAE 660  
DB 637 -----ISREQLLRFARALYPYLOAE-----LFIRMSL----- 663  
QY 661 HDQEDISPSLAQOQDKLPEPLSWRSDEDESDSDFGEQRDCYLKVSQKEHQPIITF 720  
DB 664 --DELQ-----AVIDQMLAAALVE-----QDLRLQENDTFIRPAPSSRQYVILL 705  
QY 721 LQRLGLPPLLEAYSSAIFVHNFSGVPPEYQKLYKLITRTE-----RNVVAVAESATY 776

DB 706 LARSVTQTLQRFYMAILLNAGQNALTALEENLCTVMAQRSLILHGLNAPEFPDKS-- 763  
QY 777 CLVKNVAKFKDGVKFKQKRVSVLEL 805  
DB 764 -LFRHFQIOTLLDLRVLRKDEAGKLSYHEL 791  
RESULT 7  
PLSB\_XANCP  
ID\_PLSB\_XANCP STANDARD; PRT; 886 AA.  
AC QBP3E3;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glycero1-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).  
GN PLSB OR XCC4128.  
OS Xanthomonas campestris (pv. campestris).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xanthomonas.  
OX NCBI\_TaxID=340;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 33913 / NCPPB 528;  
RX MEDLINE=22022145; PubMed=12024217;  
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
RA Camarotte G., Cannavan F., Cardozo J., Chambergos F., Ciapina L.P.,  
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,  
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.T.,  
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
RA Pereira H.A., Rossi A.J., Sena J.A.D., Silva C., de Souza R.F.,  
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
RA Setubal J.C., Kitajima J.P.;  
RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
RT host specificities."  
RL Nature 417:459-463 (2002).  
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-  
CC acyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: DE NOVO PHOSPHOLIPID BIOSYNTHESIS; FIRST STEP. MAY ALSO  
CC FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.  
CC -1- SUBCELLULAR LOCATION: Membrane-bound (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC  
CC EMBL: AE012537; AA043349.1; ALT\_INIT.  
CC HAMAP: MF\_00393; -, 1.  
CC InterPro: IPR002123; Acyltransferase.  
CC Pfam: PF01553; Acyltransferase; 1.  
CC Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;  
CC Complete proteome.  
CC SEQUENCE 886 AA; 98459 MW; A91787C8CB5341F CRC64;  
Query Match 10.6%; Score 452; DB 1; Length 886;  
Best Local Similarity 25.2%; Pred. No. 3.6e-25;  
Matches 145; Conservative 113; Mismatches 223; Indels 94; Gaps 17;  
QY 107 RLUSYVLFQERDVHKGMFATNV-----ENVLSNRVQEAIAEVAELNPDGSAQOQS 160  
DB 245 RKLQVRLTHFRIRREAVIGPDLSTRLLVDQVLAADSREATATQAKRDN-----S 296





DR InterPro; IPR002123; Acyltransferase.  
DR Pfam; PF01553; Acyltransferase; 1.  
DR SMART; SM00563; PlsC; 1.  
KW Transferase; Acyltransferase; Peroxisome; Membrane.  
FT DOMAIN 5 9 POLY-SER.  
SITE 676 678 MICROBODY TARGETING SIGNAL (POTENTIAL).  
SQ SEQUENCE 678 AA; 76869 MW; 74245A2DDDD174FF CRC64;  
  
Query Match 9.9%; Score 424; DB 1; Length 678;  
Best Local Similarity 24.0%; Pred. No. 2.8e-23;  
Matches 166; Conservative 122; Mismatches 255; Indels 148; Gaps 23;  
  
QY 133 VLNSRVQEAIAEVAELPDGSAQOQKAVKVKKAKRILOEMVAVSPAMRLTGW 192  
DB 74 VLSSEINVIKLGRE-----SLTGVDLREAESEILEENSHKLRIGAIRFFAV 124  
  
QY 193 LKLFNSPFWNIQIHKGLEVMKATETNLPLELPVHRSHDYLLTLFFLFCNKAIPY 252  
DB 125 LSKIFQIFSKVCVNEGIQKQRAVQEH-PVLLPSHRSYIDFLWLSFILSYDLPVPV 183  
  
QY 253 IASGNN-LNIPFSTLIHKLGGFFIRRLDETDPGRKDVLRALHGHIVELLR-QOQFL 310  
DB 184 IAAGMDFLGMVVSELLRMSGAFWR---TFGNGK-LYWAVSEYVKTMLRCGYAPV 237  
  
QY 311 EFLFEGTSSRSKTSKARAGLLSVVDLTSTNVIPIILIPVIGISYDRIL-EGHYNGSQL 369  
DB 238 EFLFEGTSSRAAKLTTPKFLGLNIVMEPFKREVDYTFVPITISISYDKILBESLYAYEIL 297  
  
QY 370 GPKKNESLWSVARGVIRMLKNGCVRDVFAQPSLKEYLESQKPVSAALLSLEQALL 429  
DB 298 GVPKPESTTGLK-ARRILSENFSIHVYFGDVPVSLRSLAAGRLNRNT-----YNLV 349  
  
QY 430 PAILPSR-PSDAADGDRDTSINESRNATDESRLRL--IANLAHLFTASKSCAIMGTH 486  
DB 350 PRCIQKQPEDV-----QAFVTEVAYKMLQIENLA-----LSPW 385  
  
QY 487 IVACLLLYRHRQIDLTSLVEDFVYKBEVLARDFDLGFSGNS--EDVVMHAIQLLNCV 544  
DB 386 LLVVTILLQNLQSLDFDALVEKTLWLGVTQVFGGFLWPDNKLPEEVVQSSILLHSNLA 445  
  
QY 545 TITHSRNDEPITST-----TVPSVPEL---NFSYNGVLHVFMIEAIIACSLYAV 593  
DB 446 SLV-----KQVVVLKWNSSQVVGVLVPEHIALMCSAYRNQQLNIFARPSPVALALHMT 501  
  
QY 594 --LNKRGGLGPTSTPPNLISSQBLVKKAASL-----CYLLSNEGTISLPCQTFYQVCHE 645  
DB 502 PGLREDVFCFSFLRNVSDEFILPGNTLRDPFEGCYLLCKAAMQV-----550  
  
QY 646 TVGKFIQYIGILVAEHDDQEDISPLABQQWDKLLPEPLSMRSDEDESDFGERQDCY 705  
DB 551 -AGKDI---ILT-----DK-----560  
  
QY 706 LKVSQSKHQOPIFTLQRLGLPLLEAYSAALFVNFSGVPPEPYLOKLHKYLITRTER 765  
DB 561 -----GTAVLOFLTSLFKPFVSQYLLCRYLLEEDYFGEKEYLAARFKFTQLLDQ 612  
  
QY 766 NVAVVAESATYCLVKNVAKMKDGVKFKTK 796  
DB 613 GSSQCYDALSELQKNLAFAFVRLGVVVEKK 643  
  
RESULT 11  
ID\_PLSB\_VIBVU STANDARD; PRT; 809 AA.  
AC Q8DD48;  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).  
GN PLSB OR VV1165.  
OS Vibrio vulnificus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrrio.

OX NCBI\_TaxID=672;  
RP [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=CMCP6;  
RA Rhee J.H.; Kim S.Y.; Chung S.S.; Kim J.J.; Moon Y.H.; Jeong H.;  
RA Choy H.E.;  
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-  
CC acyl-sn-glycerol 3-phosphate.  
CC -!- PATHWAY: De novo phospholipid biosynthesis; first step. May also  
CC function in the regulation of membrane biogenesis.  
CC -!- SUBCELLULAR LOCATION: Membrane-Bound (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AB016800; AAC09636.1; -;  
DR HAMAP; MF\_00393; -; 1.  
DR InterPro; IPR002123; Acyltransferase.  
DR Pfam; PF01553; Acyltransferase; 1.  
DR SMART; SM00563; PlsC; 1.  
KW Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;  
KW Complete proteome.  
SQ SEQUENCE 809 AA; 90596 MW; 0E04FF0B4980CDC CRC64;  
  
Query Match 9.3%; Score 399.5; DB 1; Length 809;  
Best Local Similarity 22.4%; Pred. No. 2.3e-23;  
Matches 180; Conservative 145; Mismatches 22; Indels 257; Gaps 38;  
  
QY 84 PSLGLRVIYINETH-----TRHRGWLARRLSYVLFTEQERDVHKG-----MFATNVT 130  
DB 169 PVVSLR---YMAQSHGTDSSIAHK---LARVARHFSRQKLAASGPDLPSSQVLFA--- 218  
  
QY 131 ENVLNSRVQPAIAEVAELPDGSAQOQKAVKVKKAKRILOEMVAVTSPAMI----- 186  
DB 219 -RLMKSPAIEQAIEE-----EAKNKNISMEKARKEADIMDEIAADPSYSLVKQGD 268  
  
QY 187 RLTCGWLLKLLFNSPFWNIQIHKGLEVMKATETNL-----PLLFLPVHRSHDYLLTLF 241  
DB 269 RLLGWL-----WN-KLYQG-LNINNAATVRRLAQDGHGHEIVYVPCRSHMDYLLSY 317  
  
QY 242 ILFCHNIKAPYIASGNLNI-----PIFSTLIHKLGGFFIRRLDETDPGRKDVLYRALLH 297  
DB 318 VLXHEGVVPHIAAGINLPPAGPIF---RRGGAFFIRSF-----KGNRLYSTIFR 367  
  
QY 298 GHIVELLRQQOFLFIELEGTSSRSKTSKARAGLLSVVDLTSTNVIPIILIPVIGISYD 357  
DB 368 EYLAELFAKGSVEYFSEGGSRRTGRLLPAKTMGLAMTIQAMRLGNRPVTLVPVYIGE 427  
  
QY 358 RIIE-GHYNGEOLCKPKKNESLWSVARGVIRMLK--NYGCVRVDFAPQPSLKEYLESQS 414  
DB 428 HVMVEVATYAKELRGKRKEKEN---AGLVRLTKLRNFGLVYGFPEIPINQVLNBSHA 483  
  
QY 415 QK-----PVSALLSLEQALLPAILPSRPSDAADGDRDTSINESRNATDESRLRLIANL 468  
DB 484 PEWTKDIDPMCA-----SRPQ-----WINP-----VVNQL 508  
  
QY 469 AEHLFTASKSCAIMSTHIVACLLLYRHRQIDLTSLVED---FFVVKERVLARDFDLG 524  
DB 509 ANKMWTHINDAAAAANALTLCATALLASQRALSXDSLHQLIECYLQLLKNVPYSKTYTVP 568  
  
QY 525 FSGNSDVMHAIOL-----LGNCVTITHSRNDEPITSTVPSVFLNFIYNSG 575  
DB 569 -SESAAALVEHAISLDKFVETDTMGDIISL---DRNQSLMT-----YRNN 612  
  
QY 576 VLHVIMEAIIACSLYAVLNKRLGGTSTTPPNLISQBLVKKAASLYLLSNEGTSILP 635









Db 211 PARQDLNFKLLAKAIARAVED-----EARSKKISHEKAQQAIAIALMBEIAANFSYEM 263  
Qy 186 IRLTGWLLKLFNSFFWNIOIHKQLEWKAATETNPLFLPVHRSHIDYLLTLTFLC 245  
Db 264 IRLTRILGTWNKLYOGINVHNA--BRVQLAHGHEIIVVPCHRSHMDYLLSYVLYH 321  
Qy 246 HNIKAPYIASGNLNI-----PIFSTLIHKLGGFIRRRRLDETDPGRDVLRYALLGHIV 301  
Db 322 QGLVPPHIAAGINLFWPAGPIF-----RLGAPFIR-----TFKGNK--LYSTVFREYLG 371  
Qy 302 ELLRQQOFLFELGTRSRGKTS CARAGLLSVVVDLTSTNVPDILIIIPVGSYDRITE 361  
Db 372 ELFSRGYSVEYFVEGGRSRTGLLDPKTGLTSLMTIQAMLRGTRPITLPIYIGYHME 431  
Qy 362 -GHVNGQLGPKKNESLWSVARVIRMLKKNYGVVDFVFAQPSLKEYLES-----QS 414  
Db 432 VGTAKELRGATKEKSLPQMLKSLK--RNLGGYVNFGEPMPLTYLNQHVPEWRES 489  
Qy 415 QKPVSALESLEQALLPA--ILPSRPSDAADSGRDTSSINESRNATDESRLRLIANLAHIL 473  
Db 490 IDPI-----EALPWLPTVNSIAAD-----LM 513  
Qy 474 FTASKCAIMSTHIVACLLYHRHQGIDLTSLVED-----FFVMKEVLAARDPDLGFSGNS 529  
Db 514 VRINNAGANAMNLCTALLASRQSLTREQLTEQLDCYLDLMRNVPYSTDSTVP-AASA 572  
Qy 530 EDVVMHAIOLGNCVTITHTSRNDEFFITPTSTVPFELNFYSNGVLHVFMIEAIACS 589  
Db 573 GELIAHALQW--NKFEVEKDTIGDIILPREQAV-----LMTYRNNIAHMLPMSMAAI 626  
Qy 590 LYAVLNKRLGGTPTPPNLISOBQLVRKAASLCYLLSNEGTSLSLPCQYQVCHETVGK 649  
Db 627 I-----TQHRISRDALQOHVEALYPMLKAEFLRWEREELASVIDALASE 672  
Qy 650 FIQVGLITVAEHDQEDISPS 670  
Db 673 MQRQGLITL--QDELHINPT 691

## RESULT 15

PLSB VIBCH STANDARD; PRT; 811 AA.  
ID PLSB VIBCH  
AC Q9KPB;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).  
GN PLSB OR VC0093.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=E1 Tor N16961 / Serotype O1;  
RX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
RT cholerae.";  
RL Nature 406:477-483(2000).  
CC -I- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-  
CC acyl-sn-glycerol 3-phosphate.  
CC -I- PATHWAY: DE NOVO PHOSPHOLIPID BIOSYNTHESIS; FIRST STEP. MAY ALSO  
CC -I- SUBCELLULAR LOCATION: Membrane-bound (by similarity).  
CC -I- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AE004100; AAF93271.1; -;  
DR EMBL; B82365; B82365.  
DR TIGR; VC0093; -;  
DR HAMAP; MF\_00393; -; 1.  
DR InterPro; IPR002123; Acyltransferase.  
DR Pfam; PF01553; Acyltransferase; 1.  
DR SMART; SM00563; PlcC; 1.  
KW Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;  
KW Complete proteome.  
SQ SEQUENCE 811 AA; 90824 MW; B076969A30BCFC8D CRC64;

Query Match 9.3%; Score 396; DB 1; Length 811;  
Best Local Similarity 23.3%; Pred. No. 4.1e-21;  
Matches 157; Conservative 119; Mismatches 217; Indels 182; Gaps 30;

Qy 47 ATLKW--KESLSMRKPPV-----GRCCYCTPQSKDFNFNPSISLGLR 89  
Db 129 ATVLWGKPKGQBRPYLQALNGPEKALAVLASGRDCL-----VRFSPV----- 174  
Qy 90 NVIYNETH-----TTRGWLARRLSYVLFIOERDVHK-----MFATNTVENVLNS 136  
Db 175 SMRYADTHGTDASTAHK--LARVARIHFSRQKLAASGNLPQRAQLFA-----RLMNS 226  
Qy 137 SRVOEAIKVAELNPDGSAQOQSKAVNKVKKAKRILEMVAIVSPAMI-----RLTGWV 192  
Db 227 PALEKAIAD-----EAKSKQIPLEKARKEHDLDEIAADFSYSLVKGDRLGLW 277  
Qy 193 LKLFNSFFWNIQIHKQLEWKAATETNL-----PLFLPVHRSHIDYLLTLTFLC 247  
Db 278 -----WN-RYQG-LNINNAATVRLAQDGHIEIVVPCHRSHMDYLLSYVLYHEG 326  
Qy 248 IKAPYIASGNLNI-----PIFSTLIHKLGGFIRRRRLDETDPGRDVLRYALLGHIVEL 303  
Db 327 MVPPHIAAGINLFWPAGPIF-----RRGAPFIRRSFKGAP-----LYSTIFREYLAEL 376  
Qy 304 LRQQOFLFELGTRSRGKTS CARAGLLSVVVDLTSTNVPDILIIIPVGSYDRITE-G 362  
Db 377 FARGSVSEYFVEGGRSRTGLLDPKTGLTSLMTIQAMLRGTRPITLPIYIGYHMEVG 436  
Qy 363 HYNGEOLGPKKNESLWSVARVIRMLK--NYGCVRVDFVFAQPSLKEYLESQSOKPVSA 420  
Db 437 TYAKELRGKKEKEN-----AGLVLRTLRLKRNFGQGVNFGFPIPLNQFLNE----- 484  
Qy 421 LLSLEOALLPAILPSRPSDAADSGRDTSSINESRNATDESRLRLIANLAHILFTASKSC 480  
Db 485 -----TVPOWTQDIDPMGE--SKPQWMTPTVKNLANRMMTHINDAAVNAATLCL 531  
Qy 481 AIMSTHIVACLLYHRHQGIDLTSLVEDFFVMKEEV--LARDPDLGF-----SGNSEDV 532  
Db 532 A-----TALLASRQ-----ALARDNLIKQVDCYLLSRNVPSYATSTLPSEAEKL 578  
Qy 533 VMHAIQL-----LGNCVTITHTSRNDEFFITPTSTVPFELNFYSNGVLHVFMIE 583  
Db 579 VEHAESLDKFFVETDMGDIISL--DRNQSIILMT-----YRNNIHLALP 623  
Qy 584 ATIASCLYAVLNKRLGGTPTPPNLISOBQLVRKAASLCYLLSNEGTSLSLPCQYQVVC 643  
Db 624 SLIAQLLIR-----QQSVLSLEKQVATVAQIYYPFLKQELFLRFAEELNDLV 669  
Qy 644 HETVKGFIQYGLITV 658  
Db 670 LRYVAELARQGLVTV 684

Search completed: January 10, 2004, 20:19:17

Job time : 29 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2004, 19:55:42 ; Search time 35 Seconds  
(without alignments)  
2275.076 Million cell updates/sec

Title: US-09-935-290-2  
Perfect score: 4280  
Sequence: 1 MDESALTGTIDVSYLPHSS.....FLPQCNRKLLLEYLSFVVL 828  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3985.5	93.1	827	A41672	glycerol-3-phospha
2	678	15.8	718	T29448	hypothetical prote
3	505	11.8	870	B82732	glycerol-3-phospha
4	462	10.8	834	F83185	glycerol-3-phospha
5	397	9.3	806	AB1015	glycerol-3-phospha
6	396	9.3	811	B82365	glycerol-3-phospha
7	395	9.2	810	D64090	glycerol-3-phospha
8	391.5	9.1	827	XURCAG	glycerol-3-phospha
9	391	9.1	825	AC0039	glycerol-3-phospha
10	390.5	9.1	827	H91256	glycerol-3-phospha
11	387.5	9.1	827	D86097	glycerol-3-phospha
12	289	6.8	775	T45238	probable transfera
13	245	5.7	789	A70868	probable plsb2 - M
14	229	5.4	621	D70762	probable transfera
15	150	3.5	287	AG2414	hypothetical prote
16	140	3.3	3144	A46068	Huntington disease
17	135.5	3.2	239	F96984	1-acyl-sn-glycerol
18	134	3.1	240	S75162	hypothetical prote
19	128.5	3.0	967	T40171	probable chromosom
20	124.5	2.9	1176	A27826	DNA-directed RNA p
21	119.5	2.8	1174	S57060	probable membrane
22	118	2.8	1258	T30252	nuclear protein SA
23	116	2.7	1262	T30524	protein phosphatas
24	115.5	2.7	3119	I49729	HD protein - mouse
25	115	2.7	3744	S46715	hypothetical prote
26	111	2.6	829	F89630	protein F59F5.7 [i
27	111	2.6	987	1 JQ0016	6-phosphofructokin
28	110.5	2.6	820	D71471	probable DNA misma
29	110	2.6	651	T15624	hypothetical prote

ALIGNMENTS

RESULT 1

A41672

glycerol-3-phosphate acyltransferase homolog - mouse  
C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 05-Nov-1999

C:Accession: A41672

R.Shin, D.H.; Paulauskis, J.D.; Moustaid, N.; Sul, H.S.

J. Biol. Chem. 266, 23834-23839, 1991

A:Title: Transcriptional regulation of p90 with sequence homology to Escherichia coli g

A:Reference number: A41672; MUID:92084678; PMID:1721057

A:Accession: A41672

A:Molecule type: mRNA

A:Residues: 1-827 <SHI>

A:Cross-references: GB:MT7003; NID:gl93366; PIDN:AAA37647.1; PID:gl93367

A:Experimental source: liver

Query Match 93.1%; Score 3985.5; DB 2; Length 827;  
Best Local Similarity 92.5%; Pred No. 7.3e-295;  
Matches 766; Conservative 32; Mismatches 29; Indels 1; Gaps 1;

QY	1	MDESALTGTIDVSYLPHSSSYSGRCRKHSTSEMGECGRPTVFRSATLKWKESLMSRKR	60
DB	1	MEESSVTGTIDVSYLPSSSYSLGRCKHSTSEDWDCGKPTFFRSATLKWKESLMSRKR	60
QY	61	PFVGRCCYSCCTPQSDWKFPNFSISLGLRNVIYINETHRHGWLARLSYLVLFQERDV	120
DB	61	PFVGRCCYSCCTPQSDWERFPNFSISLGLRNVIYINETHRHGWLARLSYLVLFQERDV	120
QY	121	HKGMPATVNTENLNSRVQEAIAEVAELNPDGSAQQSKAVNKVKKKRILQEMVAT	180
DB	121	HKGMPATVNTENLNSRVQEAIAEVAELNPDGSAQQSKAIAQKKRARKLQEMVAT	180
QY	181	VSPAMIRLTGWLKLLFNSTFNWIIQHKGLEWKAATETNLLPLFLPVHRSHIDYLLIT	240
DB	181	VSPGMRITGWLKLLFNSTFNWIIQHKGLEWKAATETNLLPLFLPVHRSHIDYLLIT	240
QY	241	FILFCHNFKAPYIAGNNLNIPFSTLIHKLGFFIRRLDETDPGRKDVLRALLHGHI	300
DB	241	FILFCHNFKAPYIAGNNLNIPFSTLIHKLGFFIRRLDETDPGRKDVLRALLHGHI	300
QY	301	VELLRQQQFLFLEBGTGRSGKTSRAGLLSVVVDTLSTNVIPDILLIPVIGISYDRIT	360
DB	301	VELLRQQQFLFLEBGTGRSGKTSRAGLLSVVVDTLSTNVIPDILLIPVIGISYDRIT	360
QY	361	EGHYNGEQLGPKPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLESQKPVSA	420
DB	361	EGHYNGEQLGPKPKKNESLWSVARGVIRMLRKNYGVVDFAPQFSLKEYLESQKPVSA	420
QY	421	LLSLEQALLPAILPSRPSDADEGDTDSINSRNATDSLRRLRIANLAHILFTASKSC	480
DB	421	PLSLEQALLPAILPSRPNDVADEHQLDSNESRNPDAEFRRLRIANLAHILFTASKSC	480

Qy 481 AINSTHIVACLLYRHRQIGDLSLVEFFVMKEEVLARDFDLGFSGNSDVMHAIQILL 540  
Db 481 AINSTHIVACLLYRHRQIGHLSLVEFFVMKEEVLARDFDLGFSGNSDVMHAIQILL 540  
Qy 541 GNCVTIHTSRNDEFFITPTSTVPSFELNPFYNGVLHVFMIMEAIIACSILYAVLNKRGGLG 600  
Db 541 GNCVTIHTSRNDEFFITPTSTVPSFELNPFYNGVLHVFMIMEAIIACSILYAVLNKRGCSG 600  
Qy 601 GPTSTPPNLIISOEOLVRKAASLCVLLSNEGTSISLPCQTFYQVCHETVKGFIQIGILTVAE 660  
Db 601 GSAGGLGNLIISOEOLVRKAASLCVLLSNEGTSISLPCQTFYQVCHETVKGFIQIGILTVAE 660  
Qy 661 HDQEDISPSGLAEQOWDKLPELWSRSDREDEDSDFGEQRDCYLKVSQSKHQQITF 720  
Db 661 QDDQEDVSPGLAEQOWDKLPE-LNWSRSDREDEDSDFGEQRDCYLKVSQSKHQQITF 719  
Qy 721 LQRLGLPLLRAYSAALIFVHNFSGVPPEPEYLOKLHKYLITRTRNVAVAESAATYCLVK 780  
Db 720 LQRLGLPLLRAYSAALIFVHNFSGVPPEPEYLOKLHKYLITRTRNVAVAESAATYCLVK 779  
Qy 781 NAVMFKDIDGVFKETKQKRVSVLESLSTFLPQCNQKLLVYLSFVVL 828  
Db 780 NAVMFKDIDGVFKETKQKRVSVLESLSTFLPQCNQKLLVYLSFVVL 827  
RESULT 2  
T29448  
hypothetical protein F08F3.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T29448  
R;Blanchard, M.; Bradshaw, H.  
submitted to the EMBL Data Library, July 1996  
A;Description: The sequence of C. elegans cosmid F08F3.  
A;Reference number: 220620  
A;Accession: T29448  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-718 <BLA>  
A;Cross-references: EMBL:U64847; PIDN:AAB04876.1; GSPDB:GN00023; CESP:F08F3.2  
A;Experimental source: strain Bristol N2; clone F08F3  
C;Genetics:  
A;Gene: CESP:F08F3.2  
A;Map position: 5  
A;Introns: 42/3; 65/2; 156/3; 279/1; 310/3; 346/3; 406/1; 516/1; 552/3; 600/3; 667/1  
Query Match 15.8%; Score 678; DB 2; Length 718;  
Best Local Similarity 27.2%; Pred. No. 3.2e-43;  
Matches 197; Conservative 148; Mismatches 264; Indels 114; Gaps 21;  
Qy 128 NVTENVLSRVQAIAEVAELNPDGSAQOQSKAVNKKKAKRILOEMVATVSPAMIR 187  
Db 78 NVEKDVLSKKRVHRVISKL-----KEQN---DEQKNRAVQFFTEISARLSKFIK 124  
Qy 188 LTGVLLKLFSFPWNTQIHKGQLEVMKATETNLPILFLPVHRSHDYLLILFILFCHN 247  
Db 125 CCSVLVYKVRRLMDKLLVCKEENEVYAEQIGIPVYPLHRSHLDYLLITWCNWHFG 184  
Qy 248 IKAPYIASGNLNIPIFTSLIHLKGGFFIRRLDPTDGRKDVLYRALLHGHVIELLRQQ 307  
Db 185 LKLPHIASGNLNLISGLWLLRATGAFIRRRVDPDDERGDQLYRALLHSYIEQVLSKD 244  
Qy 308 QFLFIIECTRSRGKTSARAGLLSVVDLTSLNIPDILIPVGLSYDRIIECHNGE 367  
Db 245 MPIEFLLEGTRSRFGKALTPKNGLISNVNVAVGFIKDCVLPVSVTYDAVVEGIFLHE 304  
Qy 368 QLGPCKNESLWSVARGVIRMLRNKNGC--VRVDPFAOPFSLKEYLESQKPSALSLSLE 425  
Db 305 LMGIPKRVESVLGVRGIFSGFSKQCGVVMHYGPIRLTEVLAT-----ITASLSN 359  
Qy -426 QALLPAILPSPDAADGDRDTSINES-----RNATDESRLRLIANLAETHILFT 475  
Db 360 HRTRPV-----RMTKLSTSFVSRELVPWHRTHSETVDDRTMIRAIGHFVVYE 406

8780/421

Qy 476 ASKSCAIMSTHIVACLLYRHRQIGDLSLVEFFVMKEEVLARDFDLGFSG---NSD 531  
Db 407 AQMCKSISPVAVVSCULLAKWKGKVSRTPERDCEWLCEKIIAEGGDVGVYQSKTKGSA 466  
Qy 532 VVMHAIQILLGNCVITHTSRNDEFFITPTSTVPSFELNPFYNGVLHVFMIMEAIIACSILY 591  
Db 467 LVKAYEKLSCSEVETI-----DE-VVSPKSHSSPITLAYNKNNSVICRPSIKSVIALTI- 519  
Qy 592 AVLNRKGLGGPTSTPPNLIISOEOLVRKAASLCVLLSNEGTSISLPCQTFYQVCHETVKGFI 651  
Db 520 -----VSRPSGK---LSIDQIVEDALSCLDWLQFEPFCRCDLSRELNVNLG--- 566  
Qy 652 QYGILTVAEHDDQEDISPSGLAEQOWDKLPELWSRSDREDEDSDFGEQRDCYLKVSQSKHQ 708  
Db 567 -----QKWS-----DPHGFRLRSEIEDDGFLDAGGALNSGTLRV 601  
Qy 709 SOSKEHQQITFLQRLGLPLLRAYSAALIFVHNFSGVPPEPEYLOKLHKYLITRTR----- 763  
Db 602 RDAKS-RETLQPPANLVRPFVQSLYLISSVVFSEKCP-EPTSDNNIIRQLCCQSLAGDI 659  
Qy 764 ERNVAUVAESAATYCLVKNVAFKDIQVFKETKQKRVSVLESLSTFLPQCNQKLLVYLSFVVL 817  
Db 660 DLFPAPLLESINSDFSNALRVLDKGLLQRTSPNSTARSGLAELIS-----NLER 713  
Qy 818 LLE 820  
Db 714 VLE 716

RESULT 3

B82732

glycerol-3-phosphate acyltransferase XF1031 [imported] - Xylella fastidiosa (strain 9a5c)  
C;Species: Xylella fastidiosa  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
C;Accession: B82732

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
Nature 406, 151-157, 2000  
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A;Reference number: A82515; PMID:20365717; PMID:10910347  
A;Note: for a complete list of authors see reference number A59328 below

A;Accession: B82732

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-870 <SIM>

A;Cross-references: GB:AE003940; GB:AE003849; NID:g9105966; PIDN:AAF83841.1; GSPDB:GN0012

A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohme

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigre

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miacca, E.C.; Miyaki, C.Y.;

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaka

M.; Tshako, M.H.; Vallada, H.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira

A;Reference number: A59328

A;Contents: annotation

C;Genetics:

A;Gene: XF1031

C;Superfamily: glycerol-3-phosphate acyltransferase

Query Match 11.8%; Score 505; DB 2; Length 870;

Best Local Similarity 22.4%; Pred. No. 6.4e-30;

Matches 197; Conservative 148; Mismatches 314; Indels 222; Gaps 26;

Qy 13 VSVLPHSSEYSVGR-CKHTSEHWGEC-----GPRPTVFRSATLKWKSLSMRKPPFVG 64

Db 64 VOYFPCVQNDGRPIIYVLEDYGLCNTLIDKACRKTLPSPILPLPGNPLOKRAYLA 123

QY 65 RCYSCYCT-----PQSWDKFNPDSIPSL-----GLRNV-----IYNETHTRHRGMLA 106  
Db 124 LSRSSNSLIPNQRGGKTHSDSLANLLQAHRIKQVIRLTLVHLPVPSIFIGRTPTDROSGWFA 183  
QY 107 -----RSLSVLFIQ 116  
Db 184 VLFSENWALVGRFRLALLVLLNGRNTIVCFAPPISVROTNLBGLPPTLRKQLQVLRH 243  
QY 117 ERDVKHGMFATNVT-----ENVLNSRVQOEAIAEVAALPDGSAQQQSKAVNKKKA 170  
Db 244 FRIRRETIVGPDLSRRLLDVNLVATEAVREAIA-----SQAKRDGDLSETWKA 294  
QY 171 KRILQEMVATVSPAMIRLTGCVLLKLFNSFFNIOIHKGQLEVMVKAATETNLPFLFVPH 230  
Db 295 QAYAMEIAADYSSPIRSADFLFSVWNRIRYAGVLH--HVDSEK-ETAPGHEVVVPSH 351  
QY 231 RSHIDYLLLTFLFCHNFKAPVIASGNLNIPFSTLHKLGFFIRRLDETTPGRKDV 290  
Db 352 RSHIDYLLLSYCLYCGGIVLPHIVAGINLNLPIVGTLLRKCKGAFIRSI-----KGNM 405  
QY 291 LYRALLHGHIVELLRQQOFLFLEGTRSRSGKTS CARAGLLSVVVDLTSTNVI PDILII 350  
Db 406 LYSIVLSEVAVQVAGVSGSYLFYFEGSRSGRLQPKGIMMTLQAFLRQPRPVLFQ 465  
QY 351 PVGISYDRIIEG-HYNGEOLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEY 409  
Db 466 PIYIGYKLIETSYLDELSGBPKKESIRLFWNIPKVLKQYQVNVNFGEPITALNDV 525  
QY 410 LESQOKPVSALESLEQALLPAILSRPSDADEGRDTSINESRNATDESRLRLIANIA 469  
Db 526 L-----AELAPEWEGQALNENKPAWLS-----NTVNHIA 555  
QY 470 EHLFTASKSCAIMSTHIVACLLVYRHOGLDLSLTVEDFFVWKEEVLARDFDLGFSG-- 527  
Db 556 RQIQTRINSADVNPINLLALLSTPHXAGEADLIAITLCKILL-----ELPYSNRV 611  
QY 528 -----NSEDDVVMHAIQILGNCTVITHTSRNDEFFITPTSTVPSEFELNFSYNGVLHVFIM 582  
Db 612 TVTPTHTPERIIAHAERQI--NILTRVHHPLGLDVLVDGNAV----LLSVFRNNVLHLFTA 665  
QY 593 EAITACSLYAVLNKRGLGFTSTPNLISQEOLVKAASLCVLLSNEGTSILPCQTFYQV 642  
Db 666 SAWVACCFKN--NR-----ISRIALIRLGVMTVPFLOAEFLPMTEDQFAQH 711  
QY 643 CHETVGKTOYXGILTVAEHDDQEDISPSLAEOQWKKLPEPLSWSDSEDESDSGEQR 702  
Db 712 IQQVIELFVREGLL--LSAGDEED-----PLTRNTSQTDE-----745  
QY 703 DCYLKVSQSKHQQFIFLQRLGLLLEA-----YSSAAIFVHNFSPGVPPEYILQKLHK 757  
Db 746 -----VFRRLRAISHLSLOQAFERYITISILVKNPGTILSASE-----LES 785  
QY 758 YLITTERNVAVIARSATVCLVKNVAKMFQIGVFKETKQK 798  
Db 786 LCQLAAQRLLSLYASTA-----PBFFDKGLFRGFTQK 817

## RESULT 4

F83185  
glycerol-3-phosphate acyltransferase PA3673 [imported] - Pseudomonas aeruginosa (strain  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83185  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bz  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: AB2950; MUID:20437337; PMID:10984043  
A:Accession: F83185  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-834 <STO>

A:Cross-references: GB:AB004787; GB:AB004091; NID:g9949835; PIDN:AAG07061.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: pldB; PA3673  
C:Superfamily: glycerol-3-phosphate acyltransferase

Query Match 10.8%; Score 462; DB 2; Length 834;  
Best Local Similarity 22.8%; Pred. No. 1.1e-26;  
Matches 171; Conservative 150; Mismatches 262; Indels 166; Gaps 27;  
QY 88 LRNVIIYNETHTRHRGMLARLSVVLFIQERDVKHGMFA-----TNVTENVLNSRVQE 141  
Db 178 LRELVEQKQGER-----TLRMVNRILRVHFRNLKTKTAVIGDLSHRRNLVGLGLRAPLVKQ 233  
QY 142 AIAEVAALNPDGSAQQQSKAVNKKAKKILQEMVATVSPAMIRLTGCVLLKLFNSFF 201  
Db 234 AISE-----ECESERISQKAEI-----ALRYANIASDFYPVIR-----FLEVILSWF 279  
QY 202 WNIQHKG-----OLEMVKAATETNLPFLFVPHRSHIDYLLTFLFCHNFKAPVIASGN 257  
Db 280 WN-KLYEGVKVNHIERVQDVAGGN-EIVVVPCHRSIDYLLSLLFRNGLTPPHIAAGI 337  
QY 258 NNTPIFSTLHKLGFFIRRLDETTPGRKDVLRALLHGHIVELLRQQOFLFLEGST 317  
Db 338 NLNMPVIGSILRGGGAFPMRESF-----KGNQLYTAVFNEYLHTLFGSGFSTFYFVEGG 391  
QY 318 RSRGKTS CARAGLLSVVVDLTSTNVI PDILII PVGISYDRIIEG-HYNGEOLGPKKNE 376  
Db 392 RSRTGRMLHPTGMLAITRLSFLRDSRRPIVFPVYIGYERVLEGRTYLGEURGATKKKE 451  
QY 377 SLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQOKPVSALESLEQALLPAILPS- 435  
Db 452 SIFDLK-VVGALQRFGQVWVNFGEPIHLQDFLRHQ-----PDW 491  
QY 436 RPSDADEGRDTSINESRNATDESRLRLIANLAEHLFTASKSCAIMSTHIVACLLYR 495  
Db 492 QDQDLGPEYRDPWLPTTNL-----LAKDVARHL-----NDAAAINPVNLVALALLST 539  
QY 496 HRQGLDLSLTVEDFFVWKEEVLARDFDLGFPS-----GNSDDVVMHAIQILL 540  
Db 540 SRQALD-----ESALARILDYLLALKRKVPSPSATLPDGGQALIEYVKS- 586  
QY 541 GNCVTITHTSRNDEFFITPTSTVPSEFELNFSYNGVLHVFIMEAIIACSLYAVLNKRGIG 600  
Db 587 -NLLAEQKDALGRILYLDQNAVLA---TYRRNVNLHVFPALPALIAFFQS--NSR--- 636  
QY 601 GPTSTPNLISQEOLVKAASLCVLLSNEGTSILPCQTFYQVCHETVGKTOYXGILTVAE 660  
Db 637 -----ISREQLLRFPALYPVLOAE-----LFIRWSL----- 663  
QY 661 HDDQEDISPSLAEOQWKKLPEPLSWSDSEDESDSGEQRDCYLKVSQSKHQQFIFTF 720  
Db 664 --DELD-----AVIDQWLAALVE-----QDLRLQENDTIFRPPSSRQYVLLIL 705  
QY 721 LQRLGLPLEAYSSAAIFVHNFSPGVPPEYILQKLHKYLIIRTE-----RNVAVYAESATY 776  
Db 706 LARSVTQTLQRFYMAIALLLNAGQNALTAELNLCVTMAQRLLSILHGLNAPEFFDKS-- 763  
QY 777 CLVKNVAKMFQDQIGVFKETKQKRVSLLEL 805  
Db 764 -LFRHFIQTLLDLRLVLRKDEAGKLSYHEL 791

## RESULT 5

AB1015  
glycerol-3-phosphate O-acyltransferase (EC 2.3.1.15) - Salmonella enterica subsp. enteri  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A:Note: this species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AB1015  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
., S.; Moule, S.; O'Gaora, P.





A; Cross-references: GB:U32758; GB:L42023; NID:gl573747; PIDN:AAC32406.1; PID:gl573755; I  
C; Superfamily: glycerol-3-phosphate acyltransferase  
C; Keywords: acyltransferase; coenzyme A; membrane protein; phospholipid biosynthesis  
C; Keywords: acyltransferase; coenzyme A; membrane protein; phospholipid biosynthesis

Query Match 9.2%; Score 395; DB 2; Length 810;  
Best Local Similarity 22.6%; Pred. No. 1.3e-21;  
Matches 166; Conservative 132; Mismatches 274; Indels 164; Gaps 28

105 LARRLSYVL---FIQERDVHKGMFATN---VTENVLNSSRVQEAIAEVAAELNPDGSAQQ 158  
 186 IAQKLARVAKMHFAKQRIATSGPRLPNRQAMFNKLQSEATFRAI-----EDEAKS 236

159 QSKAVNKVKKAKRILQEMVATVPAMIRLTGWVLKLFNSFFWNIQHKGLEMVKAAT 218  
::  
237 KNISIEKAEKYKILDELAADVSHSSLRVDRFLEWLWNKLVSNGVNONS-EVRKIAL 295

219 ETNLPFLFVPSRSHIDYLLLTFFILFCHNIKAPYIASGNLNIPIFTSLIHKLGGFTRR 278  
| : ::::| ||||| :::: : |::| :::: : |::::|  
296 EGH-EIVVPCHSRSHIDYLLLSVYLHOGIVPPHTAAGININFWPTGRMFRSNGCAFFTTP 354

279 RLDETPGRKQVLYRALLHGHIVELLRQQQFLFIUEGTRSRSGKTS CARAGLLSVVVD 338  
 355 ---TRKGNR--IVSAIPREVI.SRI.FHR.GYSVEV.FTEGCRSTGB.I.A.DKTCNMVSMY.OA 409

**339** LSTNVIPDILIIIPVGISVDRIIE-CHYNGEQLGPKKNESLWSVARGVMRLKNYGCVR **397**

: | : | : | : :: | | | | | : | : | : | : | :

**409** IOHSQTBTSIVSGMGGVELNIFVTWVAKEI DCAAYGVKKNAICI VT P VITVP D NI COGE AEE

```

398 VDFAQPSLKEYL-----ESQSQKPVSAALLSLEQALLPAILPSRPSDAADEGRDTS 448
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
457 VNECEBITI SNVI SQHRRDKWKEPNTHEFYD QWENRDS 503

```

[illegible]

```

509 FFMKEVLARFDLGFCSNSEDVVM-----HAIQLGNCVTTHTRGNDEFIT 558
      : : : | ||| : | : | | |
      : : : | ||| : | : | | |

```

559 PSTTVPSEFELNPFYNSGVLHVFI MEAIIACSLYAVLNKRGIGGPTSTPPLNISQEQLVRK 618

619 AASLCY-----LLSNEGTSIPQCQTFYQVCHETVKGKFTQYGILTVAEHDDQEDISPS 670

671 LAE--QQWDKKLPEPLSWRSDEDEDSDFGEEQRDCYLKVS-----QSKEHQ 716

691 AVRILQWSAGMUELD-----QR-IYIIVILQKQFALSRAELERESQL /33

717 FITFLQBLGLLEAYSSAAIFVHNFSGVPPEYQLKHLKYLIITRERNVAYAE--SA 774

792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

793 EICLTVKGTIERSEDL 808

RESULT 8

J. Bacteriol. 174, 5309-5316, 1992  
A; Title: Cloning and sequencing of *Escherichia coli* ubiC and purification of chaperonin 60  
A; Reference number: A42956; PMID:92355505; PMID:1644758  
A; Accession: C42956

J. Bacteriol. 174, 5309-5316, 1992  
A:Title: Cloning and sequencing of *Escherichia coli* ubiC and purification of  
A:Reference number: A42956; MUID:92355505; PMID:1644758  
A:Accession: C42956

A;Status: preliminary  
A;Molecule type: DNA  
A;Molecule type: DNA  
A;Residues: 805-827 <NIC>  
A;Note: sequence extracted from NCBI backbone (NCBIN:110475, NCBIP:110481)

R.; Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, A.; Rose, D.J.; Mau, B.; Shao, Y. Science 277, 1453-1462, 1997

A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: H65211  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A;Residues: 1-827 <BLAT>  
A;Cross-references: GB:AE000477; GB:U00096; NID:92367338; PID:  
A;Experimental source: strain K-12, substrain MG1655  
C;Citation:  
C;Comments:

A;Gene: p1eB  
A;Map position: 92 min  
C;Function:

C/Superfamily: glycerol-3-phosphate acyltransferase

Query Match	9.1%;	Score	391.5;	DB	1;	Length	827;
Best Local Similarity	24.0%;	Pred.	No. 2.6e-21;				

QY 76 DKFFNPSICLGLNVIYINETHTRHGWLARLSYVLFIQRDVKHGMFAFNVNVLN 135

Qy 136 SSRVQBAIEVAELNPDGSAQQQKAVN--KVKKAKRIIQEMVATVSPAMIRLTGWVL 193

[illegible]

DD	292	GF IWNKDIQGINVHNA--ERVQLAHGHELVVPC	349
		CHKSHMDI LLLLSI VLIEUGLVPEPHI	
QY	254	ASGNLNLI-----PIFSTLIHKLGGFFIRRR	309
		LDETPDGRKDVLYRALAHGHI VELLRQQQF	

DB 350 AAGINLNFWPAGPIF-----RKLGAFTLR-----TFKGNK---LYSTVFRFYLGLFELFSGKYS 399

QY 310 LEIFLEGRSGSGKTS CARAGLLSVVVDLTSTNVIPDILLIPVGISYDRIIE-GHYNGEQ 368

DB	400	VEYFVEGGRSRTGLLDPKTKTGLSMITQAMLRGGTRPITLPIYIGYEHVMEVGTYAKEL	459
QY	369	LGKPKNESLWSVARGVIRMLRNKYGCVRVDFAPQPSFLKEYLES-----QSOKPVSALL	422

Db	460	RGATKEKESLPQMLRGLSKL--RNLGQGYVNFGEPPMLTMYLNQHVPDWRESIDPI----	513
QY	423	SLEQALLPAILPSRPDSAADEGRDTSINESRNATDESRRRLRLIANLAEHILTASKSCAI	482

Db 514 ---EAVRPAWLTP-----TVNNIAADLMVRINNAGAA 542



QY 483 MSTHIVACLLLYRHROGIDLSTLVED----FFVWKEEVLARDFDLGFSGNSEDVVMHAIQ 538  
Db 543 NAMNLCCTALLASRQSLTREQLTEQLNCYLDLMRNVYSTDTVP--SASASELIDHALQ 601  
QY 539 LLGNCVTITHTSRNDEFFITPTTVPVSEFELNFYSGVLHVFIEMAIACSLYAVLNKRG 598  
Db 602 M--NKFEVEKOTIGDIIILPREQAV---LMTYRNNTIAHMLVLPSLMAAIV-----647  
QY 599 LGGTSTPPNLIISQOLVRKAASLYLLSNEGTSILPCQTFYQVCHETVVGKFIQYGIILTV 658  
Db 648 -----TQHRHISRDVLMHVNVLVPLMKAEFLRWRDRDELDPVIDALANEMORQGLITL 701  
QY 659 AEHDDQEDISP 669  
Db 702 --QDDELHNP 710  
RESULT 11  
DB6097  
glycerol-3-phosphate acyltransferase [imported] - Escherichia coli (strain O157:H7, sube  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
A;Accession: DB6097  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimantanta, E.; Potamouisis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: DB6097  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-827 <STO>  
A;Cross-references: GB:AB005174; NID:g12518990; PIDN:AAG59240.1; GSPDB:GN00145; UMG:P.Z56  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: plsB  
C;Superfamily: glycerol-3-phosphate acyltransferase  
Query Match 9.1%; Score 387.5; DB 2; Length 827;  
Best Local Similarity 24.1%; Pred. No. 5.2e-21;  
Matches 147; Conservative 116; Mismatches 249; Indels 99; Gaps 21;  
QY 76 DKFNPSIPSLGRNVIINETHTHRWGLARRLSYLVFIQERDVHKMFATNTVENVLN 135  
Db 182 DSFVRFS--PSVSLRRMADEHGTDKTIQAKLA-RVARMHFARQRLAAVGP-RLPARQDLFN 238  
QY 136 SSRQVEATAEVAELNPDGSAQOOSKAVN--KVKKAKRILOEMVATVSPAMIRLTGHWL 193  
Db 239 KLLASRALAKAVED-----EARSKTSISHEKAQQNAIALMEETAAANFSYEMIRLTDRI 291  
QY 194 LKLFNSFWNIQIHKGOLEMVKAATETNLPLFLFVHRSHIDYLLLTFLFCHNIKAPYI 253  
Db 292 GFTWNLVQGINVHNA--ERVQLAHGDHGLVYVPCHRSHMDYLLLSVLVHQGLVPHI 349  
QY 254 ASGNLNI-----PIPTLIHLKGGFFIRRLDETPDGRKDVLYRALLHGHIVELLRQOQF 309  
Db 350 AAGINLNPWAGPIF---RRLGAFPIRR---TFKGNK--LYSTVPREYLGELFSRGYS 399  
QY 310 LEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPTLIIPVGISYDRIIE-GHYNGEQ 368  
Db 400 VEYFVEGSRSTRGLDDPKTGMTIOAXLRGGTRPTLPIYIGYEHVNEVGTYAKEL 459  
QY 369 LGKPKKNESLWSARGVIRMLRNKNGCVRVDFAPQFSLKEYLES-----OSQKPVSA 422  
Db 460 RGATKEKESLPQMLEGLSKL--RNLGQGVNFGEPMLMTYLNQHVDPWRESIDPI-----513  
QY 423 SLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHEHILFTASKSCAI 482  
Db 514 --EAVRPWLTP-----TVNNIAADLMVRINNAGAA 542  
QY 483 MSTHIVACLLLYRHROGIDLSTLVED----FFVWKEEVLARDFDLGFSGNSEDVVMHAIQ 538  
Db 543 NAMNLCCTALLASRQSLTREQLTEQLNCYLDLMRNVYSTDTVP--SASASELIDHALQ 601

QY 539 LLGNCVTITHTSRNDEFFITPTTVPVSEFELNFYSGVLHVFIEMAIACSLYAVLNKRG 598  
Db 602 M--NKFEVEKOTIGDIIILPREQAV---LMTYRNNTIAHMLVLPSLMAAIV-----647  
QY 599 LGGTSTPPNLIISQOLVRKAASLYLLSNEGTSILPCQTFYQVCHETVVGKFIQYGIILTV 658  
Db 648 -----TQHRHISRDVLMHVNVLVPLMKAEFLRWRDRDELDPVIDALANEMORQGLITL 701  
QY 659 AEHDDQEDISP 669  
Db 702 --QDDELHNP 710  
RESULT 12  
T45238  
probable transferase [imported] - Mycobacterium leprae  
C;Species: Mycobacterium leprae  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000  
A;Accession: T45238  
R;James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, March 1999  
A;Reference number: 222949  
A;Accession: T45238  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-775 <JAM>  
A;Cross-references: EMBL:AL049913; PIDN:CAB43153.1  
A;Experimental source: cosmid B1610  
C;Genetics:  
A;Note: plsB

Query Match 6.8%; Score 289; DB 2; Length 775;  
Best Local Similarity 23.6%; Pred. No. 1.5e-13;  
Matches 102; Conservative 86; Mismatches 184; Indels 60; Gaps 10;  
QY 168 KKAKRILOEMVATVSPAMIRLTGHWLLKLFNSFWNIQIHKGOLEMVKAATETNLPLFL 227  
Db 207 EEACKMLDELATGWSRASVDLVSLGRMLSRGFPEIDYDEVQVAAMEALEAH-PAVLL 265  
QY 228 PVHRSHIDYLLLTFLFCHNIKAPYIASGNLNIPTSTLIHLKGGFFIRRLDETDPGR 287  
Db 266 FSHRSYIDGAVVPVAMQENRLPPVHVFAGINLSFGLMGLRLRRSGVIPIRRNIGDNP---322  
QY 288 KDVLRYALLHGHIVELLRQOQFLEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPTDI 347  
Db 323 ---LYKVLREYGYIVKEKPNLSWSEGTSTRTKMKLPPKGLGLTYVADAYLDGRSDDI 379  
QY 348 LIIPVGISYDRIIE-GHYNGEQLGKPKKNESLWSARGVIRMLRNKNGCVRVDFAPQFSL 406  
Db 380 LLQPVISFDQLHETAETAEVAAVARGGEKTPEGVAMLYSFIKAQGERNYGKIYVRPEAVSM 439  
QY 407 KEYLESQKPVSAALLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIA 466  
Db 440 RQYLGA---PHGALVQ-----DQDAKRLAQ 462  
QY 467 NLAHEHILFTASKSCAINSTHIVACLLLYRHROGIDLSTLVE--DFFVWKEEVLARD 520  
Db 463 KMSFEVAMRLCAPVTATALLVALLTTTRGVALTLQQLHHTLQESLDYLERKOTPVSKS 522  
QY 521 FDLGFSGNSEDVVMHAIQQL--GNCVTITHTSRNDEFFITPTTVPVSEFELNFYSGVLH 578  
Db 523 ---ALRLSRREGVRAVDALSSGHPIITRVDSGREPVMYITPGNEHAAA----FYRNSVIH 575  
QY 579 VFIMEAIACSL 590  
Db 576 AFLETSIVELAL 587

RESULT 13  
A70868  
probable plsB2 - Mycobacterium tuberculosis (strain H37Rv)  
C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C;Accession: A70868  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: A70868  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-789 <COL>  
A;Cross-references: GB:AL021246; GB:AL123456; NID:g3261507; PIDN:CAA16059.1; PID:e123760  
A;Experimental source: strain H37Rv  
C;Genetics:  
A;Gene: pIsB2

```
Query Match      5.7%; Score 245; DB 2; Length 789;
Best Local Similarity 22.0%; Pred. No. 3.4e-10;
Matches 112; Conservative 95; Mismatches 206; Indels 96; Gaps 16;

QY 116 QERDV-----HKGMPATVNTVNSRVQEAIAEVAE-----LMPDGSAAQ 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 149 QWRDVTVAEHRDPAQFVSRRAL-----LALAAEYRILGPQYKSPRLVKPEMLASA 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 159 QSKA-VNKKV---KKAKRILOENV---ATVSPAMIRLTGWLLKLFNSFFWNIQHKGO 210
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 201 RFRAGLDRIICATVEDAGKMLDELSTGWSQVSDVLSVGLRASRGDPDEF---DYDEYQ 257
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 211 LEMVKAATENLPLFLFVHRSHIDYLLATFILFCHNIKAPYIASGNLNIPFSTLIHK 270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 258 VAAARAALAAH-PAVLLFHSRYSIDGVVPMQDNRLPPVHFGGINLSFGLMGPLMR 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 271 LGGFFIRRRIDETPDGRKQVLYRAALHGHIVELLRQQOFLFLEIGHTRSRSGKTSARAG 330
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 SGMIFFTRRNIDNP-----LYKVLKEYVGVYVVEKRFNLSWTEGTSRSTGRMLPPKLG 370
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 331 LLSVVVDLTSTNVPDILLIIPVGSYDRIIE-GHYNGEQLGKPKNESLSVARGVIRML 389
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 371 LMSVADAYLDGRDDLLQGVSLCFDQLHEITEYAAARGAETPGLRWLNIFYKAQG 430
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 390 RKNYGVVRVDFAPFSLKEYLESQKPVSAALLSLEQALLPAILPSRPSDADEGRDTSI 449
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 431 ERNFGKIYVRFPEAVSNQYL-----GAPHGELTQ- 460
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 450 NESRNATDESRLRLRIANLAHLFTASKSCAIMSTHIVACLLYRHRQIDISTLVE-- 507
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 461 -----DPAAKRLALQKMSFEVAWRILOATPVTATGLVSALLITTRGTALTLDLQHLHTL 513
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 508 -----DFFVMKEEVLARDFDGFSGNSEDEVVHAIQLL--GNCVTIHTSRNDEFFITPST 561
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 514 QDSLDYLERKQSPVSTS-----ALRLRSREGVRAAADALSNGHPVTRVDSGREPVWYIAPDD 570
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 562 TVPSVFELNYSNGVLHVFIIMEAIIACSL 590
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 571 EHAAA-----PYRNSVIAHFLETSIVELAL 595
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 14  
D70762  
probable transferase - Mycobacterium tuberculosis (strain H37Rv)  
C;Species: Mycobacterium tuberculosis  
C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C;Accession: D70762  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: D70762

A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-621 <COL>  
A;Cross-references: GB:Z74020; GB:AL123456; NID:g3261584; PIDN:CAA98312.1; PID:e248544; I  
A;Experimental source: strain H37Rv  
C;Genetics:  
A;Gene: pIsB1

```
Query Match      5.4%; Score 229; DB 2; Length 621;
Best Local Similarity 19.4%; Pred. No. 3.8e-09;
Matches 134; Conservative 116; Mismatches 270; Indels 170; Gaps 24;

QY 126 ATNVTG--NVLNSSRVQEAIAEVAELNPDGSAQQQSKAVNKVKKAKRILOENVATVSP 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 29 ATDPVEVTQLDARWYDERLRALADELGRD-----PDSVRAEAAGYREMAASLDE 79
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 184 AMIR-----ITGVVLLKLFNSFFWNIQHKGOLEMWKAAATETNLPLFLPVRHSHIDYLL 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 RAVQAWRGFSRWLMRA-----YDVLVDEDOITQLR-KLDRKATLAFAFSHRSYLDGMML 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 240 TFIPLFCHNIKAPYIASGNLNIPFSTLIHKLGFFIRRRIDETPDGRKQVLYRAALHGH 299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 PEAILANRLSPALTFCGANLNFFPMGAWAKRTGAIPIRQTKDIP-----VTRFVLRAY 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 300 IVELLRQOQFLFLEIGHTRSRSGKTSARAGLLSVVVVDLTSTNVIPIDILLIIPVGSYDRI 359
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 AAQLVQNHVNLTWISIEGGRTRTGKLRPPVFGILRYITDAVEIDGDEPVVLVPTSIVVDQL 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 360 IEHG-VNGQLGPKPKNESLSVARGVIRMLR---KNYGVVRVDFAPFSLKEYLESQSQ 415
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 HEVEAMTTAYGAVKRPEDL---REFVRLARQOGERLGRAYLDGFEPLRLKRLQ--- 298
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 416 KPVSAALLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLRIANLAHLFT 475
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 -----EMRADKSGTSGEIER--IALDVEHRI-- 322
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 476 ASKSCAIMSTHIVACLLYRHRQIDISTLVEDFVMKEEVLARDFDGFSGNSEDEVVHM 535
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 323 -NRATPVTPTAVVSLALIGADR-SLSISEVLATVRPLASVIAARNWAVAGAAD----- 373
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 536 AIQLGNCVTIHTSRNDEFFITPSTVSVFELN-----FVSNVGLHV 579
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 374 ----LTRNSTRIRTLHQ-----MVASGVVSVYDAGTAVWVGIGEDQHLVAAAFVNTAIHI 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 580 FIMEAIIACSLYAVLNKRGLGSGTSTPNLISQELVRAKASLCYLLSNEGTTISLPCQTF 639
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 425 LVDRVAEALALAAATTTNG---SVSPATVRDE-----ALSRLDLKFEFLFSGRAQFE 476
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 640 YQVCHET--VGKFIQYGIILTVAEHDDQEDISPSLABEQQWKKLPEPLSMRSDDEEDSDF 697
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 477 KOLANEVLLIGSVV-----DTSKPAAADVM----- 502
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 698 GEEQRCYLVKVSQKEHQOFTFLQRLGLPALLEAYSSAA--IFVHNSGVVPPEYLOKL 755
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 503 -----RLLESAD---VLLAHVLVLPFLDAYHIVADRLAAHEDDSDSDEEGFLAECL 549
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 756 HKYLITTRERNVAVYAESATYCLVKNVKNM 785
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 550 QVGKQWELQORNTA-SAESRSMELFKTALRL 578
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 15  
AG2414  
hypothetical protein all4871 [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C;Accession: AG2414  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S. DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal  
A;Reference number: AB1807; MUID:21595285; PMID:11759840



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2004, 20:17:17 ; Search time 27 Seconds  
(without alignments)  
1297.533 Million cell updates/sec

Title: US-09-935-290-2

Perfect score: 4280

Sequence: 1 MDESALTGLTIDVSYLPHSS.....FLPCNQKQLLEYILSFVVL 828

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	462	10.8	865	US-09-252-991A-20894	Sequence 20894, A
2	438.5	10.2	862	US-09-328-352-6286	Sequence 6286, Ap
3	140	3.3	3144	US-08-246-982A-6	Sequence 6, Appl
4	140	3.3	3144	US-08-453-265-6	Sequence 6, Appl
5	140	3.3	3144	US-08-457-273B-42	Sequence 42, Appl
6	140	3.3	3144	US-08-556-419-21	Sequence 21, Appl
7	140	3.3	3144	US-09-041-886-15	Sequence 15, Appl
8	115.5	2.7	3118	US-08-457-273B-8	Sequence 8, Appl
9	115.5	2.7	3119	US-08-246-982A-16	Sequence 16, Appl
10	115.5	2.7	3119	US-08-453-265-16	Sequence 16, Appl
11	109.5	2.6	1698	US-09-315-793-12	Sequence 12, Appl
12	105.5	2.5	2188	US-09-328-352-7763	Sequence 7763, Ap
13	103	2.4	1356	US-09-770-170-6	Sequence 6, Appl
14	102	2.4	531	US-09-252-991A-24404	Sequence 24404, A
15	100.5	2.3	2930	US-09-417-822-2	Sequence 2, Appl
16	100	2.3	2108	US-09-252-991A-31502	Sequence 31502, A
17	99.5	2.3	3665	US-08-222-617A-13	Sequence 13, Appl
18	99.5	2.3	3712	US-08-222-617A-25	Sequence 25, Appl
19	98.5	2.3	383	US-08-484-575A-4	Sequence 4, Appl
20	98.5	2.3	383	US-08-477-459-4	Sequence 4, Appl
21	98.5	2.3	383	US-08-479-869-4	Sequence 4, Appl
22	98.5	2.3	383	US-08-486-414-4	Sequence 4, Appl
23	98.5	2.3	383	PCT-US94-01826A-4	Sequence 4, Appl
24	98.5	2.3	383	PCT-US94-02252A-4	Sequence 4, Appl
25	98.5	2.3	840	US-08-974-549A-8	Sequence 190, App
26	98.5	2.3	872	US-08-851-843A-8	Sequence 8, Appl
27	98.5	2.3	872	US-08-851-843A-54	Sequence 54, Appl

28	98.5	2.3	872	3	US-08-974-549A-221	Sequence 221, App
29	98.5	2.3	872	3	US-08-854-050-8	Sequence 8, Appl
30	98.5	2.3	872	3	US-08-854-050-54	Sequence 54, Appl
31	98.5	2.3	872	4	US-09-430-323-8	Sequence 8, Appl
32	98.5	2.3	872	4	US-09-430-323-54	Sequence 54, Appl
33	98.5	2.3	3712	2	US-08-222-617A-4	Sequence 4, Appl
34	98	2.3	1745	2	US-09-031-485-33	Sequence 33, Appl
35	98	2.3	1745	2	US-08-847-429A-33	Sequence 33, Appl
36	98	2.3	1745	3	US-09-065-474-33	Sequence 33, Appl
37	98	2.3	1745	4	US-09-557-034-33	Sequence 33, Appl
38	97.5	2.3	303	2	US-08-321-670-2	Sequence 2, Appl
39	97.5	2.3	303	4	US-09-215-252-3	Sequence 3, Appl
40	96.5	2.3	942	3	US-08-884-681-4	Sequence 4, Appl
41	96.5	2.3	942	3	US-09-258-643-4	Sequence 4, Appl
42	96	2.2	317	4	US-09-107-532A-3993	Sequence 3993, Ap
43	95.5	2.2	395	3	US-09-080-044-5	Sequence 5, Appl
44	95.5	2.2	395	4	US-09-531-857A-5	Sequence 5, Appl
45	95	2.2	564	2	US-08-948-569A-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1  
US-09-252-991A-20894  
; Sequence 20894, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20894  
; LENGTH: 865  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20894

Query Match	10.8%;	Score 462;	DB 4;	Length 865;
Best Local Similarity	22.8%;	Pred. No. 6.3e-39;		
Matches 171;	Conservative 150;	Mismatches 262;	Indels 166;	Gaps 27;
QY	88	LRNVIYINETHTRHGRWMLARRLSYVLFQERDVHKGWFA-----TNVTENVLNSSRYOE	141	
DB	209	LRELVEQKGHER-----TLRMVNRILRVHFNLTAVIGPDLSHRNLVKGLLRAPLVYRQ	264	
QY	142	ATAEVAALNPDGSAQQOSKAVNKVKKAKRILOEMVATVSPAMIRLTGWVLLKLFNSFF	201	
DB	265	AISE---ECESERTSQEAGEI-----ALRYANEAISDFSPVIR-----FLEVILSWF	310	
QY	202	WNIQIHKG-----QLEMYKAATETNLPFLFVHRSHDYLTLTLFILCHNFKAPIASGN	257	
DB	311	WN-KLYEGVKVNHTRVQDVAQGN-EIYVPECHRSIDYLLLSYLLFNLFGTLTPPHIAGI	368	
QY	258	NLNPIFSTLHKLGFFIRRLDETDPGRKDVLYRALLHGHIVELLRQQQFLFELRGST	317	
DB	369	NLNPVIGSLIRRGGAFFMRSP-----KGNQLYTAVNFVNLTLFSGRSTFVVEGG	422	
QY	318	RSRSGKTSFCAPAGLLSVVVDTLSTNVIPDILIIIPVIGISYDRIIEGH-YNGBQLGPKKNE	376	
DB	423	RSRTGRMLHPRTGMLAITLRSFLDRSRRPIVFVPVYIGYERVLGRTYLGELRGATKKE	482	
QY	377	SLWSVAGVTRMLRKNYGCVRVDFPAQPSLKEYLESQSKPVSAALLSLEQALLFALLPS-	435	
DB	483	SIFDLFK-VVGALKQRFQGVWVNFGEPTHLDDQFLDRHQ-----PDW	522	



; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3144 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-246-982A-6

Query Match 3.3%; Score 140; DB 1; Length 3144;

Best Local Similarity 18.7%; Pred. No. 0.00048;

Matches 160; Conservative 130; Mismatches 319; Indels 248; Gaps 40;

Qy 57 SRKPPFVGRCCYCTPQSWDKFPNPSIPSLGRNVYINETHRHGWLARRLSYVLFQ 116

Db 1350 SSVRPGLYHYCF-MAPYT---HFTQALADASLRNMVQA-EQENDTSGWFD-----VLQ 1397

Qy 117 ERDVHKGMFATNTVEN-----VLNRRVQGAIAEVAALNPDGSAQOQSKAVNKKVKA 170

Db 1398 KVTQLKNTLTSVTKRADKNAIHNHILRPELVIKALK-----QYTTTCVOLQKQV 1450

Qy 171 KRILQEMVATVSPAMIRLTGWVL-----LKLNSFFWNIQIHKQLEMVKAATET 220

Db 1451 LDLLAQLV-----QLRVNYCLLSDQVFIGFVLKQFE-----YIEVGQFR---ESEA 1494

Qy 221 NLP-----LLFLPVHRSHIDL--LLTFILFCHNIKAPYIASGNL---NIPFSTLIHK 270

Db 1495 IIPNIFPVLVLSYVYHSHQIIGIPKIIQLCDGI---MASGRKAVTHAIPALQPIVHD 1550

Qy 271 LGGFFIRRLDETDPGRKDVLYRALLHGHIVELLRQOQFLEIFLEGTRSGKTSARAG 330

Db 1551 L---FVLGRTNKADAGKELETQKEVVVSMRLRLIYHQVLEMF-----KLNEDKWK-----RLSR 1610

Qy 331 LLSVVVDLTSTNVPDILIIIPVGISYDRIIEGHYNGEQLGPKPKNESLWSVARGVIRMLR 390

Db 1591 -----ILVLQ-----QCH-----KLNEDKWK-----RLSR 1610

Qy 391 KNYGCVRVDFAPQPSLKEYLESQKPVSALESQALLPAILSPRPSDADEGRDTSIN 450

Db 1611 Q-----IADIILPMLAKQMHDSHEALGVLTNLFILAPSSL--RPVDMLLRSMFVTPN 1663

Qy 451 ESRNATDES-----RRRLIANLAELHILFASKSCAISMTHIVACLLILYHRQGLD 502

Db 1664 TMASVSTVOLMISGILAILRVLISQSTEDIVLSRIQELUS-FSPYLISCTVINLRDGDST 1722

Qy 503 STLVE-----DFFVMKEEVLARDFDLGFG--NSEDVNMHAIQLLGNCTVITHTSRNDEFF 556

Db 1723 STLEHSGKQIKNLPEETFSR-FLQLVGILLIEDIVTKLK-----VEMSEQQHTFY 1774

Qy 557 ITPSTTVSPVELNPFYNGVLHVFI--MEAIACSLYAVLNKRGLGGTSTPPNLSIQEQ 614

Db 1775 CQ-----ELGTLMLCLHIFKSGMFRITAAATRLFRSDGCGG-----SFYTLDS 1819

Qy 615 LVRAASLCYLLSNEGTSLSLQTFYQVCHETVKGFIQYGLITVAEHDDQEDISPSLAEQ 674

Db 1820 LNLKARS--ITTHPALVLLWCQ-----ILLVNHDTY-----R 1851

Qy 675 QWDKKLPEP-----LSWRSDDEDESDFGEEQDCYLK-----VSQS 711

Db 1852 WVAEVQQTPKHSLSTSLSPQSGEEDSDLAALGCMCNREIVRGALILFCDYVCQN 1911

Qy 712 KEHQOFTFL-----QRLLG-----PLJEAYSSA-----AIFV-----HNFSGFVP 747

Db 1912 LHDSEHLTWLVNHIODLISLSHEPPVDFTISAVHRNSAASGLFIQAIQSCENLSPTM 1971

Qy 748 EPEVLQKHLKYLITERNVAVYAESATYCLVKNVAKMFKDIGVFKETQKRVSVLESS 807

Db 1972 LKLTQLQLEGHLQSOGAVLTLVD-RLLCPTFRVLARMVDILACRRVEMLLAANLQSSM 2030

Qy 808 TFLPQCNRQKLEYILS 824

Db 2031 AOLPMELNRQIEYLQS 2047

RESULT 4

US-08-453-265-6

; Sequence 6, Application US/08453265

; Patent No. 5693757

; GENERAL INFORMATION:

; APPLICANT: MacDonald, Marcy E.

; APPLICANT: Ambrose, Christine M.

; APPLICANT: Duyao, Mabel P.

; APPLICANT: Gusella, James F.

; TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox

; STREET: 1100 New York Avenue

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/08/453,265

; FILING DATE: 30-MAY-1995

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Ludwig, Steven R.

; REGISTRATION NUMBER: 36,203

; REFERENCE/DOCKET NUMBER: 0609.3880003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3144 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-453-265-6

Query Match

Best Local Similarity 18.7%; Pred. No. 0.00048;

Matches 160; Conservative 130; Mismatches 319; Indels 248; Gaps 40;

Qy 57 SRKPPFVGRCCYCTPQSWDKFPNPSIPSLGRNVYINETHRHGWLARRLSYVLFQ 116

Db 1350 SSVRPGLYHYCF-MAPYT---HFTQALADASLRNMVQA-EQENDTSGWFD-----VLQ 1397

Qy 117 ERDVHKGMFATNTVEN-----VLNRRVQGAIAEVAALNPDGSAQOQSKAVNKKVKA 170

Db 1398 KVTQLKNTLTSVTKRADKNAIHNHILRPELVIKALK-----QYTTTCVOLQKQV 1450

Qy 171 KRILQEMVATVSPAMIRLTGWVL-----LKLNSFFWNIQIHKQLEMVKAATET 220

Db 1451 LDLLAQLV-----QLRVNYCLLSDQVFIGFVLKQFE-----YIEVGQFR---ESEA 1494

Qy 221 NLP-----LLFLPVHRSHIDL--LLTFILFCHNIKAPYIASGNL---NIPFSTLIHK 270

Db 1495 IIPNIFPVLVLSYVYHSHQIIGIPKIIQLCDGI---MASGRKAVTHAIPALQPIVHD 1550

Qy 271 LGGFFIRRLDETDPGRKDVLYRALLHGHIVELLRQOQFLEIFLEGTRSGKTSARAG 330

Db 1551 L---FVLGRTNKADAGKELETQKEVVVSMRLRLIYHQVLEMF-----KLNEDKWK-----RLSR 1610

Qy 331 LLSVVVDLTSTNVPDILIIIPVGISYDRIIEGHYNGEQLGPKPKNESLWSVARGVIRMLR 390

Db 1591 -----ILVLQ-----QCH-----KLNEDKWK-----RLSR 1610

Qy 391 KNYGCVRVDFAPQPSLKEYLESQKPVSALESQALLPAILSPRPSDADEGRDTSIN 450

Db 1611 Q-----IADIILPMLAKQMHDSHEALGVLTNLFILAPSSL--RPVDMLLRSMFVTPN 1663







Db 1946 TTLKKTQCLGHIHSQSGAVLTYVDRLLGTSSRALARMV-DTLACRRVEMLLAANLQS 2004

Qy 806 SSTFLPQCNRQKLEIYI 822

Db 2005 SMAQLPEEELNRIQEH 2021

RESULT 9

US-08-246-982A-16  
; Sequence 16, Application US/08246982A  
; Patent No. 5686288

GENERAL INFORMATION:

; APPLICANT: MacDonald, Marcy E.  
; APPLICANT: Ambrose, Christine M.  
; APPLICANT: Duyao, Mabel P.  
; APPLICANT: Gusella, James F.  
; TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/246.982A  
; FILING DATE: May 20, 1994  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

; NAME: Goldstein, Jorge, A.  
; REGISTRATION NUMBER: 29,021  
; REFERENCE/DOCKET NUMBER: 0609.3880002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3119 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-246-982A-16

Query Match 2.7%; Score 115.5; DB 1; Length 3119;  
Best Local Similarity 17.4%; Pred. No. 0.18;  
Matches 167; Conservative 145; Mismatches 343; Indels 303; Gaps 43;

Qy 12 DVSYLPHSEYSGVKHSTSEWEGCGRPVTVFRSATLKWESLMRSKRPVVGRC----- 66

Db 1221 DVLKATHANYKVTLDNSTKFG-GF---LRSALDVLQLELATLQDQIGKVEVL 1274

Qy 67 -----CYS-----CTPQSWDFKFNFSIPSL--GLRNVIYINETHRHGWLARLSY 111

Db 1275 GYLKSCFSREPMATVVCVQLLKTFLGTNLASQFDGLSS---NPSKSQCR---AQLLG- 1326

Qy 112 VLFQIERDVHKGMF-----ATNVNVLNSSRVQEAIAEVAEELNPDG-----SA 156

Db 1327 -----SSSVRPLHYHCFWAPYTHFTQALASLNMVQAE--QERDASGDFVLQKVA 1379

Qy 157 QOQKAVNKVKKA-KRIQEMVATVSPAMIRLTGVLVLLKLFNSPFNNIQHKGLEMYK 215

Db 1380 QLKTLNLTSTVKNRADKNAHNRILFEPLVIK-----ALKQYTT-TTSVQLQKQVLDLLA 1433

Qy 216 AATENLPLFLPVHRSHIDVLLTF----- 241

Db 1434 QLVQLRVNVCLLDSQVDFIGVLFKQFIEVQGFREBAIIPNIPFFVLVLSYERYHSKQ 1493

Qy 242 -----ILFCHNTKAPYIASGNNL--NIPIFTSLIHKLGFFIRRLDFTPOGRKDV 291

Db 1494 IIGIPKIIQLCDGI---MASGRKAVTHAIPALQPIVHDL---FVLRGTNKADAGKELE 1546

Qy 292 YRALLGHIVELLRQOQFLFIEGTRSRSKTSCARAGLLSVVVDLTSTNVIPDILLIP 351

Db 1547 QKEVVVSMJLRJLIQYHQVLEMF-----ILVLQ 1573

Qy 352 VGISYDRIIEGHYNGEQLGKPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLE 411

Db 1574 -----QCH-----KENEKWK-----RLSRQ-----VADIILPLAKQOMH 1604

Qy 412 SQSQKPVSALESLEQALLPAILPSPDAADGSDTTSINESRNATDESRL----- 461

Db 1605 IDSHEALGVNLTFEILAPSSL--RPVDM--LRSMFTPTSTMASVSTVQLWISGILAIL 1660

Qy 462 RRLTANLAEHILFTASKSCAI---MSTHIVACLLLYRHROG---IDL-----STIV 506

Db 1661 RVLISQSTEDIVL-----CRIQELSFPHLLSCPVINLRGGGNGVNLGECSEKQKSLP 1715

Qy 507 EDFE-----VMKEEVLARDFDLGFSGNSEPDVVMHAIQLLGNLCVVTITTSRNDEP-F 556

Db 1716 EDTESRELLQVLGILLEDIVTKQKVDMSQOQHTFYCOELGTLLMC--LIHIFSGMPPR 1773

Qy 557 ITPSTT---VPSVPELNFYNGVNLHVFIME-----AIIACSLYAVLN----- 595

Db 1774 ITAAATRLFTSDGCEGSFYTLESINARVSRMVPHPALVLLWCQILLINHTHRMWAEV 1833

Qy 596 -----KRGLGPTSTPPNLISOEOLVRKAASL--CYL-LSNEGTSISLPCQTFYQVCHETV 647

Db 1834 QQTPKRHSLSCTKSLNPKSGEEDSGSAQAGMCNREIVRRGAILILFCDYVCQNLHDS- 1892

Qy 648 GKFIQYGLITVAEHDDQEDISPLAEQWDQCKLPELSWSRDEDESDDFGEEQDRCVLK 707

Db 1893 -----EHLTWLVINHIQD-----LIS 1908

Qy 708 VSQSKHQQTITLQF---LIGPLLEAYSSAAIFVHNFSFGVPPEVYLQKLHLYITRTE 764

Db 1909 LSHPEPPQDFISAIHRNSAASGLFIQIQSRC---ENLSTPTTLKKTLCLEGIHLSQSG 1965

Qy 765 RNVAIYAESATYCLVKNVAMFKDIGVKETKQKRVSVLELSSTFLPCNCRQKLEIYI 822

Db 1966 AVLTLYVDRLLGTTPFRALARMV-DTLACRRVEMLLAANLQSSMAQLPEEELNRIQEH 2022

RESULT 10

US-08-453-265-16  
; Sequence 16, Application US/08453265  
; Patent No. 5693757

GENERAL INFORMATION:

; APPLICANT: MacDonald, Marcy E.  
; APPLICANT: Ambrose, Christine M.  
; APPLICANT: Duyao, Mabel P.  
; APPLICANT: Gusella, James F.  
; TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/453.265  
; FILING DATE: 30-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:



NAME: Ludwig, Steven R.  
REGISTRATION NUMBER: 36,203  
REFERENCE/DOCKET NUMBER: 0609.3880003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3119 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-453-265-16

Query Match 2.7%; Score 115.5; DB 1; Length 3119;  
Best Local Similarity 17.4%; Pred. No. 0.18; Mismatches 343; Indels 303; Gaps 43;  
Matches 167; Conservative 145

QY 12 DVSYLPHSSYSYVGRCKHTSEWGECCGFRPTVFRSATLKWKESLMGRKRPFGVRC-----66  
DB 1221 DVLKATHANYKVTLDLQNSTEKG--GF-----LRSALDVLSQLLELATLQDICKVEEVL 1274  
QY 67 -----CVS-----CTPOSWKFNPSPSL--GLRNVIYNETHRRHGLMARLSY 111  
DB 1275 GYLKSCFSREPMATVCVQLLKTFLGTLNASQFDGLSS-----NPSKQCR---AORLG- 1326  
QY 112 VLFQERDVHKGME-----ATNVTENLVNLSRVQEAIAEVAELNPDG-----SA 156  
DB 1327 -----SSSVRPGLYHYCFMAYPTHFTQALADASLRNVQAE--OERDASGFVLOKWSA 1379  
QY 157 QOSKAVNKVKKA-KRILQEMVATSPAMIRLTGWLLKLFNFSFWNIQIHKQGLEMVK 215  
DB 1380 QLKTNLTSTVTKRADKNAHNLPEPIVK-----ALKQYTT-TTSVQLQKQVLDLLA 1433  
QY 216 AATETNLPPLFLPVHRSHIDYLLTF-----241  
DB 1434 QLVQLRVNYCLDSQDFVIGFLVKQFEYIEVGQFSEAIIPNIFFLVLLSYERYHSKQ 1493  
QY 242 -----ILCHNIKAPIYASGNL---NIPFSTLIHKLGGFFIRRLDETDPGRKQVL 291  
DB 1494 IIGPKIIQLCDGI---MASGRKAVTHAIPALQPIVHDL---FVLGRTKNADAGELET 1546  
QY 292 YRALLGHIVELLRQQQFLBIFLEGTRSRGKTSARAGLLSVVVDLTSTNVPIDILIIIP 351  
DB 1547 QKEVVSMRLRLQYHQVLEMF-----ILVLQ 1573  
QY 352 VGISYDRIIEGHYNGEOLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLE 411  
DB 1574 -----QCH-----KENEDKW-----RLSRQ-----VADIIPLMLAKQOMH 1604  
QY 412 SOSQKPSALLSLEQALLPAILSRPSDADEGRDTSINESRNATDESIR-----461  
DB 1605 IDSHEALGVNLTLEIAPSSL--RPVDMIL--LRSMFITEPTMASVSTVQLWISGILAIL 1660  
QY 462 RRIIANLAHEILTASKSCAI-----NSTHIVACLLYRHRQ-----IDL-----STLV 506  
DB 1661 RVLIQSOTEDIVL-----CRIQELSPHLLSCPVINLRGGGNTVLCESGKQKSLP 1715  
QY 507 EDFP-----VMKEEVLARDFDLGFSGNSEDVVMHAIQLLGNVCVTIHTSRNDEF-F 556  
DB 1716 EDTFSRLLQVGLLEDIVTKQKVDMSQQHTFYCOELGTLLMC--LHIHFKSGWFR 1773  
QY 557 ITPSTT---VPSVFELNFYNGVLHVIME-----AIIACSLYAVLN-----595  
DB 1774 ITAAATRLFTSDCEGSFYTLESNARVSRMVPTHPALVLLWCQILLNHTDHRWAEV 1833  
QY 596 -----KRGGLGPTSTPNLISQELVRKAASL--CYL--LSNEGFTISLPCQTFYQVCHETV 647  
DB 1834 QQTPKHSLSCTSLNPQKSGEEDSGSAQLGNCNREIYVRGAILFCDYVCQNLDHS- 1892  
QY 648 GKFIQVILTVAEHDDQEDISPSLAEOQWDKULPEPLSWRSDEDESDSGEQRDCYLK 707  
DB 1893 -----EHLTWLVNHIQD-----LIS 1908

QY 708 VQSKEHQOIFITFLOR---LLGLLEAYSSAAIFVHNFSGPVPEPEYLOKLHKYLITRTE 764  
DB 1909 LSHEPPVQDFISAIHRNSAASGLFQAIQSRC---ENLSTPTTLTKTLQCLSGIHLSQSG 1965  
QY 765 RNVAVYAESATYCLVKNAVNFKDIGVFKETKQKRVLSLELSTFLPQCNQKLLBYI 822  
DB 1966 AVTLYVDRLLGTTPFRALARMV-DTLACRRVEMLLAANLQSSMAQLPEBELNRIOEHL 2022

RESULT 11  
US-09-315-793-12  
Sequence 12, Application US/09315793  
Patent No. 6221597  
GENERAL INFORMATION:  
APPLICANT: Roberts, Christopher J.  
TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL  
TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATION  
FILE OF INVENTION: DRUGS  
FILE REFERENCE: 9301-048  
CURRENT APPLICATION NUMBER: US/09/315,793  
NUMBER OF SEQ ID NOS: 62  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 12  
LENGTH: 1698  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-315-793-12

Query Match 2.6%; Score 109.5; DB 3; Length 1698;  
Best Local Similarity 18.2%; Pred. No. 0.26;  
Matches 163; Conservative 137; Mismatches 289; Indels 309; Gaps 44;

QY 75 WDKFNPSPISGLRNVIYNETHRRHGLMARLSYVLFQERDVHKGMAFNVTENVL 134  
DB 706 WKSLKDPRHQILGVKNLKTNR-----FIPSKFTESALVYTFVEE--- 745  
QY 135 NSRVQEAIAEVAELNPDGSAQKQSKAVNKVKKAKRILQEMVATSPAMIRLTGWVL- 193  
DB 746 -----EDISEKLSVL--DLLMTQDSDSNLRRPLELILGELFDDQNPFFYLTVSKWILS 797  
QY 194 -----LKLNFNSFWNI---QIHKQGLEMVKAAATETNLPFLPVHRSHIDYLLTFI 242  
DB 798 ILNSGSASRLFYLTNTNLKVNLEKRLD-----ERDDLMLTYEFQ 840  
QY 243 LFCHNIKAPIYASGNLNIPIFSTLIHKLGGFFIRRLDETDPGRKQVLYRALLHGHIVE 302  
DB 841 MLAYVLK-----TNGRTRKVFSTELTSIKSTIKWNEDVS-----TYKSL--LVT 885  
QY 303 LLRQQQFLBIFLEGTRSRGKTSARAGLLSVVVDLTSTNVPIDILIIIPVGISYDRIIEG 362  
DB 886 LMR---FLNI-KSNTHAKSIRSALI---LLDILLDGTQCN-FKDIVIFLLQWSSKYIABE 937  
QY 363 HYNGEOLGPKKNESLWSVARGVIRMLRKN-----YGCVRVDF-----400  
DB 938 GIEPELIA-----VSLDIVSKVLRSLSHNGIKLIDFDDNAAHLKY-----IDFLVTSVN 988  
QY 401 -AQPSLKEYLSQSKQPSVALLSLEQALLPAILSRPSDADEGRDTSINESRNATDES 459  
DB 989 MKSPLIIVAYVKLLSE---SIVYPENSIFRMILP-----1019  
QY 460 LRRRLIANLAHEI--LFTASK-----SCAITSHTIVACL-----LLYRHRQ--IDL 503  
DB 1020 -----LSASLVQCVRQLFLEKEGGYQPIALLLGLLEELLEISHGYLVEEREGVFSG 1075  
QY 504 TLVEPDFVMKEEVLARDFDLGFSGNSE-----DVVMHAI--QLLGNVCVTIHTSR- 551  
DB 1076 NLKGDPI---QSVSNVFS--SDSSNEESKIQERDVILQSFQVISCCLDIWYWAHNISC 1131  
QY 552 ---NDEFTTPTSTVPSEFELNFYNGVLH-VFIMEAI-----IACSLYAVL- 594  
DB 1132 KSNDDSSL--DATNHSYKFKFRSKKLTLETLFLEPLELLENLISRSNTTTLVHVLD 1189

QY 595 -NKRGGLGPTSTPPNLI-----SQEOLVRKAASLCYLLSNEG 630  
 Db 1190 GNK-----PAITIPHLXYGVIRYNRTASVKFSNRDGRSSTTKLTGKPSMLKRLSGES 1244  
 QY 631 TISL-----PCQTFYOVCHETGKFIQYGLTVAEHDDQEDISPSLAQOWDKK 679  
 Db 1245 IIAELFNVDVSENSAMEEFY-----GDELLPFREVATNLYSDVLSIL-----K 1291  
 QY 680 LPEPLSWRSDDEDESDGEBQR-----DCYLK----- 707  
 Db 1292 LVALISGVSK-----TQGEQKVRREISDVFFKYLPNAFINFTNLYRGHPDSPKDLFV 1347  
 QY 708 -----VSQKEHQQTITFLORLLGPLELAYSSAAI--FVHNFSVPVPEYLQK 754  
 Db 1348 VWRVQYIVNDQIGDKFNTLTIATVNOCLTPYIKPKSEKTIPIGVVLELAADVSHLGSVKV 1407  
 QY 755 LHKYLIT--RTERNAVVAESATYCLVKNVAKMFKDIGVFKETQKQVSV--VLELSS 807  
 Db 1408 SWRLLIABLFQNDKLSVIGSDQW-----EKIYEWISIYPENKSKITLNDLLEIGS 1459  
 RESULT 12  
 US-09-328-352-7763  
 ; Sequence 7763, Application US/09328352  
 ; Patent No. 6562958  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 ; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GTC99-03PA  
 ; CURRENT APPLICATION NUMBER: US/09/328.352  
 ; CURRENT FILING DATE: 1999-06-04  
 ; NUMBER OF SEQ ID NOS: 8252  
 ; SEQ ID NO 7763  
 ; LENGTH: 2188  
 ; TYPE: PRT  
 ; ORGANISM: Acinetobacter baumannii  
 US-09-328-352-7763

Query Match 2.5%; Score 105.5; DB 4; Length 2188;  
 Best Local Similarity 18.8%; Pred. No. 1.1;  
 Matches 113; Conservative 98; Mismatches 220; Indels 169; Gaps 27;  
 QY 272 GGFFIRRLDTPDGRKDV-LYRALLHGHIVELLRQQQFLBIFLEGTRSRSGKTSK- 326  
 Db 1115 GKFTDRELDKFAEQLRGVNOYQVDPDLIDYQDMDLQRYAEG-RTTKGKLSLWIE 1173  
 QY 327 ARAGLLSVVDTLSTNVI-----PDILIPVGI-SYDRIIEGHYNGSQLGPKKNE 376  
 Db 1174 AKYGRDSIYQISIKGAYGFAQRGGENPDVVAEMFGYSGDALIRDLNSP---SPQKI 1230  
 QY 377 SLMSVARGVIRMLKRYCCVRVDFAPQPSLKEYLESQSKPVA-LISLEQALLPAILPS 435  
 Db 1231 DELTDARMAVQY-----SEFFDQOSIITEAVEAALHNDVARMLSAEMAALNGLL-- 1279  
 QY 436 RPSDAADGRTSINE-SRNATDESRLRLIANAEHLFTASKSCAITSMTIHVACILLY 494  
 Db 1280 -----GRKSALNEAKTVAQDIVQKQIKDIRPHV----- 1309  
 QY 495 RHRQIDILSTLVEDFFVNMKEEV-----LARDFDLGFSGNSEDVVMHAIQLL-----G 541  
 Db 1310 RAQDDARLGRMANEAFKGETVEAARHKNQLVQFYATKYSYDAKQDTQKHLDLVKVFG 1369  
 QY 542 NCVTITHSRNDEP-FITPSTTVSVFELNYSNGVLHVFIMEAIIACSLVAVLNKRGGLG 600  
 Db 1370 NNEKL---SKNRDFDVTAARGILGKYDLGRESTNYEHOELIRKIDYPTTVAETQNIQ-- 1424  
 QY 601 GPTSTPPN-----LISQELVRKAASLCYLLSNEGTLISLCQTFYOVCHETVCKFTQYG 654  
 Db 1425 ---ALPENQNTRELTFQFNVAAMAAVETLMHRSKEN-----KIWHTTNEAF----- 1467  
 QY 655 ILTVAEHDDQEDISPSLAQOWDKKLPPLSWRSDDEDESDGEBQRDCYLKVSQS-KE 713

Db 1468 -----EREQVREELIQOTGGKKSVKEIKQOTLLGRDKTA-----ELKAKFMEIGASAKR 1515  
 QY 714 HQQFITFLQ-----RLLGPLELAYSSAAI-FVHNFSVPVPEYLQKHLKYLITRTE 764  
 Db 1516 VDQVVTWDDGANGKFKFTYLINPMQDA-----LAKTRIEKAK 1552  
 QY 765 RNVAVYAESATYCLVKNVAKMFKDIGVFKETQKQVSVLELSS-TFLPQCNRQKLLLEYIL 823  
 Db 1553 -----MLKDVDVDFEGFGKLDNSK---IAAPELNNFTFV---GKQSLHLHAIL 1593  
 RESULT 13  
 US-09-770-170-6  
 ; Sequence 6, Application US/09770170  
 ; Patent No. 6319679  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McKnight, Steven  
 ; APPLICANT: Gardner, Kevin  
 ; APPLICANT: Harper, Shannon  
 ; APPLICANT: Rutter, Jared  
 ; APPLICANT: Michnoff, Carolyn  
 ; APPLICANT: Amezcua, Carlos  
 ; TITLE OF INVENTION: PAS Kinase  
 ; FILE REFERENCE: UTSD:0761  
 ; CURRENT APPLICATION NUMBER: US/09/770.170  
 ; CURRENT FILING DATE: 2001-01-26  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: Patentin version 3.0  
 ; SEQ ID NO 6  
 ; LENGTH: 1356  
 ; TYPE: PRT  
 ; ORGANISM: Yeast  
 US-09-770-170-6

Query Match 2.4%; Score 103; DB 4; Length 1356;  
 Best Local Similarity 19.0%; Pred. No. 0.84;  
 Matches 89; Conservative 82; Mismatches 173; Indels 124; Gaps 21;  
 QY 297 HGHIVELLRQQQFLBIFLEGTRSRSGKTSKARAGLLSVVDTLSTNV---IPDILIIIPVG 353  
 Db 324 HGPSMDLLNEQ-----RANLKS-----LLNLNETLEKNTSERASDLHMSL- 365  
 QY 354 ISYDRIIEGHYNGSQLGPKKNSLWSVARGVIRMLKRYCCVRVDFAPQPSLKEYLESQ 413  
 Db 366 -----FNLKMLGDPKKNNSERDKRTEKLLKILLD-----SLAEFF--EHNFI 409  
 QY 414 SQKPVSAALLSLEQALLPAILPSRPSDAADGRTSINESRNATDESRLRLIA---NLAE 470  
 Db 410 GDNPIADTDELKEI-----DEFTGSGDTTATDIRPQDYGRILRTFTSTKNSAP 460  
 QY 471 HILFTASKS--CAIMSTHIVACILLYRHRQIDILSTLVEDE-----FVMKEEVLRDFFD 522  
 Db 461 QAITFTCSQEDPWFQFRAANDLACLVFGISQNAIRALITMDLIHTDSRNFVLHKLSTREG 520  
 QY 523 LGFSGNSEDVVMHAIQLLGCNVTITHTSRNDEPFTPTTTPSVS-VFELNFSY---NGVLH 578  
 Db 521 MVFTG-----EIIIG-----IVPETLSSSKVWVWASFWAKRKNGLL- 555  
 QY 579 VFIMEAIIACSLVAVLNKRGGLGPTSTPPNLIHQEOLVRKAASLCYLLSNEGTLISLCQOT 638  
 Db 556 VCVFEKVPYCDVVDVLLNLDLDFGA-----ENIVDKCELLSDGPT---LSSSSTLSLP--- 603  
 QY 639 FYQVCHETVCKFTQYGLTVAEHDDQEDISPSLAQOWDKKLPPLSWRSDDEDESDG 698  
 Db 604 --KMASSPTGSKLEY-----SLERKILEKSYTKPTSTENRNGDENQLDG 645  
 QY 699 BEQRDCYLKVS--OSKEHQF-----ITFLQRLGLPLLEAYSSAAIF 738  
 Db 646 DSHSEPSLSSSPVRSKRSKVFANDIKDVKSISQSLAKLMDDVRNGVVF 693  
 RESULT 14

US-09-252-991A-24404  
; Sequence 24404, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24404  
; LENGTH: 531  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24404

Query Match 2.4%; Score 102; DB 4; Length 531;  
Best Local Similarity 21.9%; Pred. No. 0.2;  
Matches 61; Conservative 41; Mismatches 106; Indels 70; Gaps 11;  
QY 179 ATVSPAMIRLTGWLKLFNSFFWNQIHKGQLEWVKAATETNLPLFLPVHRSHIDYLL 238  
DB 309 ADVTLRQRLTRWLRLCAALFEVRV-----SGEAPROPMLVANHVSWTDIPL 359  
QY 239 LTFILFCHNFKAP--YIAGNNLNIPFSTLIHKLGGFFIRRLDETTPDGRKDVLYRALL 296  
DB 360 LGAL-----APLTLFKAQVRAWPLAGLAELKAGTLYIRR--GSGDSR-----LI 402  
QY 297 HGHIVELLRQQQLEIFLEFTRSGKTSRAGLLSVVDLTSTNVIPDILLIPVGISY 356  
DB 403 NORLAEQLHGRNLLIFPEGTITNGESLRTFHGLMASALEA-----GVAVQPVVAISY 455  
QY 357 DR----IIEGHYNGE-----OLGKPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSL- 406  
DB 456 RRGVDPDTQAFPGIDDDLLSHLG-----RLLGERGSHVHQLLEPIPSQ 499  
QY 407 ---KEYLESQKPVSAALSLEQALLPAILPSRPSDAA 441  
DB 500 GLDRAELARQAQAV-----RLALFGTAAPTQTRRAA 531

RESULT 15  
US-09-417-822-2  
; Sequence 2, Application US/09417822  
; Patent No. 6344549  
; GENERAL INFORMATION:  
; APPLICANT: Keegan, Kathy  
; TITLE OF INVENTION: ATR-2  
; FILE REFERENCE: 27866/35633  
; CURRENT APPLICATION NUMBER: US/09/417,822  
; CURRENT FILING DATE: 1999-10-14  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: Patentin ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 2930  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-417-822-2

Query Match 2.3%; Score 100.5; DB 4; Length 2930;  
Best Local Similarity 19.5%; Pred. No. 6;  
Matches 137; Conservative 102; Mismatches 254; Indels 209; Gaps 34;  
QY 218 TETNLPLELPVHRSHIDYLLLTFLFCHNFKAPYIAGNNLNIPFSTLIHKLGGFFIR 277  
DB 14 SETYALFLSLF-----SFHKFCKGLLANTVEDVNICIQACSSL-HALS----- 56  
QY 278 RLDETDPGRKDVLYRALLHGHIVELLRQQQLEIFLEFTRSGKTSRAGLLSVV 335

Db 57 ---SSLPD---DILORC-----VDVCR-----VOLVHSGTRIRQAQFGL-----LKSIP 94  
QY 336 VD-TLSTNVIPDILLIPVG-----ISVDRIIEGHYNGEQLGKPK 373  
Db 95 LDVVLNNNHTEIQEISLALRSHMSKAPNTTHPQDFSDVISF--LYG--NSHRTGKON 150  
QY 374 KNESLWSVARGVIRMLRKNYGCVRV-----YSCORLDRQDSTIPRNLKTDVLMQWAIWEAAQFTVLSK 409  
Db 151 WLERLF-----YSCORLDRQDSTIPRNLKTDVLMQWAIWEAAQFTVLSK 197  
QY 410 LESQSQKPVSAALSLEQALLPAILPSRPSDADEGRDTSINESRNATDSLRRLRLIANLA 469  
Db 198 LRTPLGRAQDTTQTIE-GIIRSLAAHTLNPQDVQSWTTADNDEGHGNNQLRLVLLQLYL 256  
QY 470 EH---ILFTASKSC--AINSTHIVACLLLYRHRQGI-DLSTLVEDDFVMKEEVLARDFDL 523  
Db 257 ENLEKLMYNAYEGCANALTSPPKVIIRTFYTNRTQCDWLTRIR-LSIMRVGLLA----- 310  
QY 524 GFSGNSDDVMEAIQLLGNCVITITHTSRNDEFFITPSTTPSVFELN-----FYSN 574  
Db 311 ---GQPAVTVRHGFDDLTEMKT-TLSLQNELEVTIMVVEALCELHCPEAIOGIAVMS 366  
QY 575 GVL-----HVFIMEAIIAC-SLY--AVLNKRGLGGPTS 604  
Db 367 SIVGKNLLWINSVAQQAEGREFKASVEYQEHLCAMTGVDCICISSFDKSVLTLANAGRNA 426  
QY 605 TPNLISQQLVYRKAASLCYLLSNEGTISLPQCTFYOVCHETVKGFIQYIGILTVAEHDDQ 664  
Db 427 SPKHSNGE-----SRKTVLSKPTDSSPEVINYLGNKACEYI----- 464  
QY 665 EDISPLAQWQDKKLP-PLSWRSEDEDESDFGEEQRDCYLKVSQSKEHQOQIFLOR 723  
Db 465 -SIADWAVQEQWNAIHDLKKSTSTSLNLKADFN-----YKLSLSESGKFEVCTEQ 517  
QY 724 L-LGP-----LLEAVSSAAIFVHNFSGPV--PEPEYLOKLHKYLIIRTRNVA-----VY 770  
Db 518 LELLPGENINLAGGSKKIDMKLLPNMSPDRELQKSIQVQLLRSSVCLATALNPTE 577  
QY 771 ABSATYCLVKNVAKFKDIGVFKEYKQKRVSVLELSSTLPQ 812  
Db 578 QDQKQOSITENVVK-----YLKQTSRIAIGPLRLSTLTVSQ 613

Search completed: January 10, 2004, 20:22:33  
Job time : 32 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2004, 20:20:58 ; Search time 58 Seconds  
(without alignments)  
2876.725 Million cell updates/sec

Title: US-09-935-290-2

Perfect score: 4280

Sequence: 1 MDESALTGTIDVSYLPHSS.....FLPQCNRQKLEYILSFVWL 828

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:\*
- 1: /cgn2\_6/ptodata/1/pubpaa/PCT\_US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4280	100.0	828	11	US-09-935-290-2
2	4279	100.0	828	12	US-10-288-252-6
3	936.5	23.3	801	12	US-10-288-252-7
4	928.5	21.7	775	12	US-10-353-690-32
5	678	15.8	718	12	US-10-369-493-6322
6	505	11.8	870	12	US-10-369-493-17551
7	500	11.7	754	12	US-10-369-493-9385
8	490	11.4	754	12	US-10-369-493-13643
9	486	11.4	754	12	US-10-369-493-9240
10	452	10.6	772	12	US-10-369-493-16014
11	452	10.6	792	12	US-10-369-493-15639
12	452	10.6	795	12	US-10-369-493-16367
13	407.5	9.5	698	12	US-10-369-493-19420
14	245	5.7	789	10	US-09-712-363-244
15	128	3.0	1095	12	US-10-094-749-1905

16	119.5	2.8	1174	15	US-10-205-841-40	Sequence 40, Appl
17	117.5	2.7	195	9	US-09-817-910-11	Sequence 11, Appl
18	117.5	2.7	195	9	US-09-798-029-27	Sequence 27, Appl
19	117.5	2.7	195	9	US-09-887-389-4	Sequence 4, Appl
20	117.5	2.7	195	10	US-09-924-358-46	Sequence 46, Appl
21	117.5	2.7	195	10	US-09-882-872-4	Sequence 4, Appl
22	114	2.7	714	12	US-10-369-493-10138	Sequence 10138, A
23	111	2.6	987	12	US-10-369-493-21998	Sequence 21998, A
24	108	2.5	788	12	US-10-369-493-11178	Sequence 11178, A
25	108	2.5	1420	12	US-10-310-154-387	Sequence 387, App
26	108	2.5	3529	12	US-10-165-216-8	Sequence 8, Appl
27	108	2.5	3657	12	US-10-165-216-10	Sequence 10, Appl
28	106	2.5	532	12	US-10-104-047-2555	Sequence 2555, Ap
29	103	2.4	432	15	US-10-220-380-4	Sequence 4, Appl
30	103	2.4	1356	15	US-10-059-963-6	Sequence 6, Appl
31	103	2.4	3031	12	US-10-165-216-6	Sequence 6, Appl
32	103	2.4	3521	12	US-10-165-216-2	Sequence 2, Appl
33	102.5	2.4	1847	12	US-10-032-585-7538	Sequence 7538, Ap
34	102	2.4	454	11	US-09-963-339-2	Sequence 51, Appl
35	102	2.4	454	12	US-10-145-586-51	Sequence 22419, A
36	102	2.4	700	12	US-10-369-493-22419	Sequence 21903, A
37	102	2.4	805	12	US-10-369-493-433	Sequence 433, App
38	102	2.4	4746	12	US-10-369-493-22294	Sequence 22294, A
39	101	2.4	1661	12	US-10-369-493-19598	Sequence 19598, A
40	100.5	2.3	615	12	US-10-369-493-19598	Sequence 2, Appl
41	100.5	2.3	2930	11	US-09-957-837A-2	Sequence 3180, Ap
42	100	2.3	463	12	US-10-094-749-3180	Sequence 5843, Ap
43	99	2.3	550	10	US-09-738-626-5843	Sequence 20, Appl
44	99	2.3	4351	12	US-09-970-944-20	Sequence 8, Appl
45	98.5	2.3	872	10	US-09-843-676-8	

ALIGNMENTS

RESULT 1

- US-09-935-290-2
- Sequence 2, Application US/09935290
- Publication No. US20030044948A1
- GENERAL INFORMATION:
- APPLICANT: Kapeller-Libermann, Rosana
- TITLE OF INVENTION: 56919, A NOVEL HUMAN ACYLTRANSFERASE AND USES THEREOF
- FILE REFERENCE: MNI-186
- CURRENT APPLICATION NUMBER: US/09/935,290
- CURRENT FILING DATE: 2001-08-21
- PRIOR APPLICATION NUMBER: 60/226,509
- PRIOR FILING DATE: 2000-08-21
- NUMBER OF SEQ ID NOS: 3
- SOFTWARE: PatentIn Ver. 2.0
- SEQ ID NO 2
- LENGTH: 828
- TYPE: PRT
- ORGANISM: Homo sapiens
- US-09-935-290-2

Query Match	100.0%	Score	4280	DB	11	Length	828
Best Local Similarity	100.0%	Pred. No.	0				
Matches	828	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MDESALTGTIDVSYLPHSSEYSVGRCKHTSEEWGCGFRPTVFRSATLKWESLMSRKR	60				
Db	1	MDESALTGTIDVSYLPHSSEYSVGRCKHTSEEWGCGFRPTVFRSATLKWESLMSRKR	60				
Qy	61	PFVGRCCYCTPQSDWDFNPSLGLRNVIIYNETHRHGWLARLSVLFQERDV	120				
Db	61	PFVGRCCYCTPQSDWDFNPSLGLRNVIIYNETHRHGWLARLSVLFQERDV	120				
Qy	121	HKGWFATNVNTVNLSSRVQEAIAEVAELNPDGSAQQSKAVNKKKAKRIIQEWAT	180				
Db	121	HKGWFATNVNTVNLSSRVQEAIAEVAELNPDGSAQQSKAVNKKKAKRIIQEWAT	180				
Qy	181	VSPAMIRLTGVLTKLFNSFFWNIQIHKGQLEWKAATETNLPILFLPVHRSHDYLLLT	240				
Db	181	VSPAMIRLTGVLTKLFNSFFWNIQIHKGQLEWKAATETNLPILFLPVHRSHDYLLLT	240				

Db 181 VSPAMIRLTGWVLLKLFNSFFWNIQIHKGQLEWVKAATNTLPLLPVHRSHIDYLLLT 240  
Qy 241 FILFCHNIKAPYIAGNNLNIPFSTLIHKLGGFFIRRRDETPDGRKDVLYRALLHGI 300  
Db 241 FILFCHNIKAPYIAGNNLNIPFSTLIHKLGGFFIRRRDETPDGRKDVLYRALLHGI 300  
Qy 301 VELLROQOFLFLEI FLEGRSRSGKTS CARAGLLSVVDTLSTNVIPDILIPVGISYDRII 360  
Db 301 VELLROQOFLFLEI FLEGRSRSGKTS CARAGLLSVVDTLSTNVIPDILIPVGISYDRII 360  
Qy 361 EGHNGEOLGPKKNESLWSVARGVIRMLRKNYGCVRVDFPAQPFSLKEYLESQKPVSA 420  
Db 361 EGHNGEOLGPKKNESLWSVARGVIRMLRKNYGCVRVDFPAQPFSLKEYLESQKPVSA 420  
Qy 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
Db 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
Qy 481 AIMSTHIVACLLYRHRQGDISTLVEDFFVMKEEVLARDFDLGFSGNSDEVVMAIQLL 540  
Db 481 AIMSTHIVACLLYRHRQGDISTLVEDFFVMKEEVLARDFDLGFSGNSDEVVMAIQLL 540  
Qy 541 GNCVTITHTSRNDEFFITPTSTVPSPFELNFGVNLHVIMEAIIACSLYAVLNKRGIG 600  
Db 541 GNCVTITHTSRNDEFFITPTSTVPSPFELNFGVNLHVIMEAIIACSLYAVLNKRGIG 600  
Qy 601 GPTSTPNLISQOLVRKAASCLYLLSNEGTSILPCQTFYQVCHETVKGFIQVGIITVAE 660  
Db 601 GPTSTPNLISQOLVRKAASCLYLLSNEGTSILPCQTFYQVCHETVKGFIQVGIITVAE 660  
Qy 661 HDQEDISPSPSLAEQWQDKKLPEPLSWRSDEDESDSDFGEQRDCYLKVSQKEHQOIFITF 720  
Db 661 HDQEDISPSPSLAEQWQDKKLPEPLSWRSDEDESDSDFGEQRDCYLKVSQKEHQOIFITF 720  
Qy 721 LQRLGLPLLEYASAAIFVNFSGPVPEPEYLOKLHKYLIITRERNVAVVAESATYCLVK 780  
Db 721 LQRLGLPLLEYASAAIFVNFSGPVPEPEYLOKLHKYLIITRERNVAVVAESATYCLVK 780  
Qy 781 NAVMFKDGVFKETKQKRVSVLELSSTLPQCNROKLLBYILSFVVL 828  
Db 781 NAVMFKDGVFKETKQKRVSVLELSSTLPQCNROKLLBYILSFVVL 828

RESULT 2

US-10-288-252-6  
; Sequence 6, Application US/10288252  
; Publication No. US20030143686A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: LAL, Preeti G.  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: YUE, Henry  
; APPLICANT: BURFORD, Neil  
; APPLICANT: GANDHI, Aneena R.  
; APPLICANT: WARREN, Bridget A.  
; APPLICANT: YAO, Monique G.  
; APPLICANT: TRIBOULEY, Catherine M.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: LEE, Ernestine A.  
; APPLICANT: HAFALIA, April J.A.  
; APPLICANT: LU, Yan  
; APPLICANT: GRIFFIN, Jennifer A.  
; APPLICANT: SANJANWALA, Madhu S.  
; APPLICANT: DING, Li  
; TITLE OF INVENTION: TRANSFERASES  
; FILE REFERENCE: PI-0241 USA  
; CURRENT APPLICATION NUMBER: US/10/288,252  
; CURRENT FILING DATE: 2002-11-04  
; PRIOR APPLICATION NUMBER: PCT US01/30424  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 60/252,819  
; PRIOR FILING DATE: 2000-11-21  
; PRIOR APPLICATION NUMBER: US 60/249,639

; PRIOR FILING DATE: 2000-11-16  
; PRIOR APPLICATION NUMBER: US 60/247,931  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: US 60/246,001  
; PRIOR FILING DATE: 2000-11-03  
; PRIOR APPLICATION NUMBER: US 60/244,025  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: US 60/238,481  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/236,523  
; PRIOR FILING DATE: 2000-09-29  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PERL Program  
; SEQ ID NO 6  
; LENGTH: 828  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030143686A1 1517230CDI  
US-10-288-252-6

Query Match 100.0%; Score 4279; DB 12; Length 828;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 827; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDSEALTCTIDVSYLPHSSEYSGRCCKHTSSEWGECPRTVPSRATLKWESLSMRKR 60  
Db 1 MDSEALTCTIDVSYLPHSSEYSGRCCKHTSSEWGECPRTVPSRATLKWESLSMRKR 60  
Qy 61 PFVGRCCYCTQSQWMDKFFNPSPISGLRNVIYINETHTRHGWLARLSYVLFQERDV 120  
Db 61 PFVGRCCYCTQSQWMDKFFNPSPISGLRNVIYINETHTRHGWLARLSYVLFQERDV 120  
Qy 121 HKGMFATNTENVLNSRRVQEAIAEVAELNPDGSAQOQSKAVNKVKKAKILQEMVAT 180  
Db 121 HKGMFATNTENVLNSRRVQEAIAEVAELNPDGSAQOQSKAVNKVKKAKILQEMVAT 180  
Qy 181 VSPAMIRLTGWVLLKLFNSFFWNIQIHKGQLEWVKAATNTLPLLPVHRSHIDYLLLT 240  
Db 181 VSPAMIRLTGWVLLKLFNSFFWNIQIHKGQLEWVKAATNTLPLLPVHRSHIDYLLLT 240  
Qy 241 FILFCHNIKAPYIAGNNLNIPFSTLIHKLGGFFIRRRDETPDGRKDVLYRALLHGI 300  
Db 241 FILFCHNIKAPYIAGNNLNIPFSTLIHKLGGFFIRRRDETPDGRKDVLYRALLHGI 300  
Qy 301 VELLROQOFLFLEI FLEGRSRSGKTS CARAGLLSVVDTLSTNVIPDILIPVGISYDRII 360  
Db 301 VELLROQOFLFLEI FLEGRSRSGKTS CARAGLLSVVDTLSTNVIPDILIPVGISYDRII 360  
Qy 361 EGHNGEOLGPKKNESLWSVARGVIRMLRKNYGCVRVDFPAQPFSLKEYLESQKPVSA 420  
Db 361 EGHNGEOLGPKKNESLWSVARGVIRMLRKNYGCVRVDFPAQPFSLKEYLESQKPVSA 420  
Qy 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
Db 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
Qy 481 AIMSTHIVACLLYRHRQGDISTLVEDFFVMKEEVLARDFDLGFSGNSDEVVMAIQLL 540  
Db 481 AIMSTHIVACLLYRHRQGDISTLVEDFFVMKEEVLARDFDLGFSGNSDEVVMAIQLL 540  
Qy 541 GNCVTITHTSRNDEFFITPTSTVPSPFELNFGVNLHVIMEAIIACSLYAVLNKRGIG 600  
Db 541 GNCVTITHTSRNDEFFITPTSTVPSPFELNFGVNLHVIMEAIIACSLYAVLNKRGIG 600  
Qy 601 GPTSTPNLISQOLVRKAASCLYLLSNEGTSILPCQTFYQVCHETVKGFIQVGIITVAE 660  
Db 601 GPTSTPNLISQOLVRKAASCLYLLSNEGTSILPCQTFYQVCHETVKGFIQVGIITVAE 660  
Qy 661 HDQEDISPSPSLAEQWQDKKLPEPLSWRSDEDESDSDFGEQRDCYLKVSQKEHQOIFITF 720  
Db 661 HDQEDISPSPSLAEQWQDKKLPEPLSWRSDEDESDSDFGEQRDCYLKVSQKEHQOIFITF 720

QY 721 LQRLGPLEAYSSAAIFVHNFSPVPEYLOKLHKYLIITERNVAVYAESATYCLVK 780  
Db 721 LQRLGPLEAYSSAAIFVHNFSPVPEYLOKLHKYLIITERNVAVYAESATYCLVK 780  
QY 781 NAVKMKDGVPEKTKQKRVSVLESLSTFLPQCNRQKLEYSFVL 828  
Db 781 NAVKMKDGVPEKTKQKRVSVLESLSTFLPQCNRQKLEYSFVL 828

RESULT 3

US-10-288-252-7  
; Sequence 7, Application US/10288252  
; Publication No. US20030143686A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: LAL, Preeti G.  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: YUE, Henry  
; APPLICANT: BURFORD, Neil  
; APPLICANT: GANDHI, Ameena R.  
; APPLICANT: WARREN, Bridget A.  
; APPLICANT: YAO, Monique G.  
; APPLICANT: TRIBOULEY, Catherine M.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: LEE, Ernestine A.  
; APPLICANT: HAFALIA, April J.A.  
; APPLICANT: LU, Yan  
; APPLICANT: GRIFFIN, Jennifer A.  
; APPLICANT: SANJANWALA, Madhu S.  
; APPLICANT: DING, Li  
; TITLE OF INVENTION: TRANSFERASES  
; CURRENT APPLICATION NUMBER: US/10/288,252  
; FILE REFERENCE: PI-0241 USA  
; CURRENT FILING DATE: 2002-11-04  
; PRIOR APPLICATION NUMBER: PCT US01/30424  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 60/252,819  
; PRIOR FILING DATE: 2000-11-21  
; PRIOR APPLICATION NUMBER: US 60/249,639  
; PRIOR FILING DATE: 2000-11-16  
; PRIOR APPLICATION NUMBER: US 60/247,931  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: US 60/246,001  
; PRIOR FILING DATE: 2000-11-03  
; PRIOR APPLICATION NUMBER: US 60/244,025  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: US 60/238,481  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/236,523  
; PRIOR FILING DATE: 2000-09-29  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PERL Program  
; SEQ ID NO 7  
; LENGTH: 801  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20030143686A1 5665262CD1  
US-10-288-252-7

Query Match 23.3%; Score 996.5; DB 12; Length 801;  
Best Local Similarity 33.0%; Pred. No. 1.1e-90;  
Matches 273; Conservative 146; Mismatches 348; Indels 61; Gaps 24;  
QY 25 CRCKTSEWGECPRTVFRSA---TLKWESLSMRKRPVGRCCYSCCTPQSWDKFFNP 81  
Db 7 GRCC-TQPRSPSGREASLSSGFGMKLEAVTPFLGKVRPFGVRCCTCTPKSWESLPHR 65  
QY 82 SIPSGLRNVIYNTHRRHGWLARLSYVLFQIORDVHKMGFATNVENLVNLSRRVQE 141  
Db 66 SITDLGFCNVILVKEBTRFRGWLVRRLCYFLWSLEQHIPP---CQDVPQKIMESTGVON 122

QY 142 AIAEVAELNPDGSAQQSKAVNKVKKAKRILOEMVATVSPAMIRLTGWLLKLFNSFF 201  
Db 123 LLSGRV---PGGTGEGQ--VPDLVKEVQRILGHIOAPPEPFLVRFSWALLRFLNCLF 176  
QY 202 WNIQIHGQLEWVAAATETNLPLFLPVHRSHIDYLLLTFLFCHNINKAPVIASNNLNI 261  
Db 177 INVQLHRGOMKVMQAAAGLPLVLLSTHKTLLDGIILPFLPMLLSQGLGVLVAVMSRACS 236  
QY 262 PIPSTLJHKLGGFFIRRLDETDPGRKDVLYRALLHGHIVELLRQOQFLEIFLEGTSRS 321  
Db 237 PALRALLRKLGLFLPPEASLSLDSSEGLARAVVQAVIEQLLVSGQPLLIIFLEPPGAL 296  
QY 322 G-KTSCARAGLLSVVVDLTSTNVIPDILIIIPVGISYDRIIEGHYNGQLGPKPKNESLWS 380  
Db 297 GPRLSALGQAVGVVQVQVGI VPDALLVPVAVTYDLPDAPCDIDHASAPL---GLWT 353  
QY 381 VARGVIRMLRKNYG-----CVRVDFAPFSLKEVLESQSQKPVSAALLSLEQALLPAILP- 434  
Db 354 GALAVLRSLMSRWGCSHRICSRVHLAQPFSLQEIYIVS-ARSCWGRQTLQQLQPIVLGQ 412  
QY 435 -SPESDAADGRDPTINESRNATDESRLRLIANLAEHILFTASKSCAIMSTHIVACILL 493  
Db 413 CTAVPDTEKQEWTPITGPILLALKEE-DQLLVRLSCHVLSASVGSASVMSATMAITLL 471  
QY 494 YRHQGIIDLSTLVBDFVFMKEEVLARDFDLGFGNSDVEDVVMHAIQLLNCNCVTIHTSRND 553  
Db 472 FKQKGVFLSQLGGEFSWLTETILLRGDFVGSQLSRLQHSLSLLRAHVALLRI-RQG 530  
QY 554 EFFITPTTVPSPFELNFGVNLHVFIMEAIIACSLYAVLNKRGLOGGPTSTPPN----- 608  
Db 531 DLLVVPQPG-PGLTHLAQLSAELLPVFLSEAVGACAVRGLLAGR-----VPPQGPWEL 582  
QY 609 ----LISOEOLVRKAASLCVLLSNEGTSILPCOTFYQVCHETVCKFYQYGLITVAEHDDQ 664  
Db 583 QGILLTSQNELYRQIILLMHLLPODLLLKPCCSSYCYCOEVLRLIQCGLL-VAEETPG 641  
QY 665 EDISPSLAEQQWDKKLPEPLSWR--SDEEDSDS-DFGEQRDCYLKVSQSKEHQOFTIFL 721  
Db 642 SRPACDTGRQELSRK----LLWPSGDFDTSDSDDFGEAD-GRYFRLSQSHCHPDPFLFL 696  
QY 722 QRLGPLEAYSSAAIFVHNFSPVPEPE--YLOKLHKYL-ITERNVAVYAESATYCL 778  
Db 697 CRLLSPLLKAFAPAAAFIR--QGOLPDTELGYTQLFQFLOATAQEEGI---FECADPKL 751  
QY 779 VKNVAKMKDGVPEKTKQKRVSVLESLSTFLPQCNRQKLEYSFVL 826  
Db 752 AISAVWTFRDLGVLLQQTSPAGPRHLSPFTFASLDNQEKLQFIQRFI 799

RESULT 4

US-10-353-690-32  
; Sequence 32, Application US/10353690  
; Publication No. US20030215840A1  
; GENERAL INFORMATION:  
; APPLICANT: Logan, Thomas Joseph  
; APPLICANT: Chun, Miyoung  
; APPLICANT: Galvin, Katherine M.  
; APPLICANT: Healy, Aileen  
; APPLICANT: Acton, Susan L.  
; APPLICANT: Donoghue, Mary  
; APPLICANT: Stagliano, Nancy  
; APPLICANT: Perodino, Jacqueline  
; APPLICANT: Rodrigue-Way, Amelie  
; TITLE OF INVENTION: Methods and compositions for treating  
; TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,  
; TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,  
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,  
; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,  
; TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,  
; TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,  
; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2254, 3474,  
; TITLE OF INVENTION: 9792, 15400, 1452 or 6585 molecules





```
Db 407 AQMCSIQPVAVVSCLLAKMGKVSRTFFERDCWLCEKIIAEGDVVGQSKTKGSA 466
Qy 532 VMHAIQILGNVCVITHTSRNDEFFITPSTTVPSVFNFSYNGVLHVFIMEAIIACSLY 591
Db 467 LVKYAFKLESCEVT-----DE-YVSPKESHSSFITLAYKNKSVICFSIKSVIALTI- 519
Qy 582 AVLNRKGLGGTSTPPNLSIQBQVRKAASLCYLLSNGTISLPQCTFYQVCHETVGKFI 651
Db 520 -----VSRPSGTK---LSIDQIVEDALSCLDWLQFEBFCPCDSRLRELHNVLG--- 566
Qy 652 OXGILTVAEHDDQEDISPSLAEQOMDKKLPEPLS--WSDERDED--SPFGEEQRCYLVK 708
Db 567 -----QKENS-----DIHGFLKSEIEDDGLDAGGALNSGTURV 601
Qy 709 SOSKEHQQFITFLQRLGLLGLLEAYSAAFVHNFSGPVPEYLOKLHKYLITRT----- 763
Db 602 RDAKS-RETLOQFANLVRPFVQSLVLISSFVVEKCP-EPTSDNNIIRQLCQOSLAGDI 659
Qy 764 ERNVAIVAESATYCLVKNAVAMFKDIGVFKETKQKRV-----VLELSSTFLPOCNRQK 817
Db 660 DLPFAPLESINSDSFKNALRVLDKGLLQSTPNSTARSGNSRLAELIS-----NLER 713
Qy 818 LLE 820
Db 714 VLE 716

RESULT 6
US-10-369-493-17551
; Sequence 17551, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17551
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-369-493-17551

Query Match 11.8%; Score 505; DB 12; Length 870;
Best Local Similarity 22.4%; Pred. No. 5.1e-41;
Matches 197; Conservative 148; Mismatches 314; Indels 222; Gaps 26;

Qy 13 VSYLPHSSEYSVGR-CKHTSEBWGEC-----GPRPTVFRSATLKWKESLMSRKPFGV 64
Db 64 VOYEHCVQYNDGRPIIVLEDEYGLCNTLLIDKACRKTLPSPILPNCNPLQRRAYLA 123
Qy 65 RCCVST-----PQSWDKFNPSPSL-----GLRNV-----IYNETHRRHGWLA 106
Db 124 LSRSSNSLIPNQRGGKTHSDSLANLLQAHRIRDTLDVHVPVSIFIGRTPDQSGWFA 183
Qy 107 -----RRLSYLVFIQ 116
Db 184 VLFSENWALVGRFRLLAVLLNGRNTIIVCFAPPIISVROTLINEGLPPTLRLKQRLRAH 243
Qy 117 ERDVHKGMFATNVT-----ENVLNSSRVQEAIAEVAELPDGSAQOOSKAVNVKKKA 170
Db 244 FRRIRETVIGDPLSTRLLLVNDVLAATEAVRAIA-----SOAKRDGDTLSETWRKA 294
Qy 171 KRILQEMVATVSPAMIRLTGWLLKLKLFNSFFWNIQIHKQGLQEMVKAATETNLPLFLPVH 230
```

```
Db 295 QAYAWETAAADSSPVIRSADFLFSGHVNRIYAGVLIH--HVDSPK-ETAPGHEVVYVPSH 351
Qy 231 RSHDYLILLTILFCHNIKAPYIASGNLNIPIFTLHLKLGFFIRRLDETDPGRKDV 290
Db 352 RSHDYLILLSYCLTQCQGVILPHIVAGINLNIPIVGTLLRKCGAFFIRSI-----KGNM 405
Qy 291 LYRALLHGHIVELLARQQQFLBIFLEGTRSRSGKTSACARAGLLSVVVDTLSTNVPIDILII 350
Db 406 LYSIVLSEYVAQLVAGVSLVLEYFIEGGRSRTGRLLPQKGGIMMTLQAFLLQPRPVLQ 465
Qy 351 PVGISYDRITEG-HYNGEQLGKPKKNSLWSVARGVIRMLRKNYGCVRVDPQAQPSLKEY 409
Db 466 PIYIGYKLTGTSYLDLSGEPKKEISIRLFWNIPIKVLKQKYGVVNVFGEPIALNDV 525
Qy 410 LESQSPVPSALLSLEQALLPAILPSRPSDAADGRDTSINESRNATDESRLRLIANLA 469
Db 526 L-----AELAPWEQOALNENKPAWLS-----NTVNHLA 555
Qy 470 EHILFTASKCAIMSTHIVACLLYHRQGIIDLSLTVEDFFVMKEEVLARDFDLGFSG-- 527
Db 556 RQIQTRINSAADVNPINLLALALLSTPKHAMEADLIAQITLCKKILL-----ELPYSNRV 611
Qy 528 -----NSEDVVMHAIQLLGNVCVITHTSRNDEFFITPSTTVPSVFNFSYNGVLHVFIM 582
Db 612 TVTPHTPERIIAHAEQI--NILTRVHHPGLGDVLRVDGDNV-----LLSYFRNNVLHLFTA 665
Qy 583 EAIACSILYAVLNKRGGLGGTSTPPNLSIQBQVRKAASLCYLLSNGTISLPQCTFYQV 642
Db 666 SAWVACCFKN--NRR-----ISRLLIRLGVGYFPFLOABLFLPWTEDQFAQH 711
Qy 643 CHETVGKFIQVGIITVAEHDDQEDISPSLAEQOMDKKLPEPLSWRSDDEDSDFGEBQR 702
Db 712 IQQVIELFVREGLL-LSAGDEED-----PLTRNTQGTDE----- 745
Qy 703 DCYLVKSQSKHEHQFITFLQRLGLLGLLEA-----YSSAAIFVHNFSGPVPEYLOKLHK 757
Db 746 -----VFLRAISHSLQAQAFERYITISILVKNKGGTSLASE-----LES 785
Qy 758 YLITRTERNVAVVAESATYCLVKNAVAMFKDIGVFKETKQK 798
Db 786 LCQLAAQRLSLLYASTA-----PEFFDKGLFRGFIQK 817

RESULT 7
US-10-369-493-9385
; Sequence 9385, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9385
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-369-493-9385

Query Match 11.7%; Score 500; DB 12; Length 754;
Best Local Similarity 22.1%; Pred. No. 1.3e-40;
Matches 194; Conservative 152; Mismatches 308; Indels 222; Gaps 26;

Qy 18 HSSYSVGR-CKHTSEBWGEC-----GPRPTVFRSATLKWKESLMSRKPFGVGRCCYS 69
```

Db 3 HSYQVNDGRPIIVLEVDYGCNTLILDKACRTKLPSPILPDPNPLQKRAYLALSRSS 62  
Qy 70 CT-----PQWDKFFNPSISL-----GLRV-----IYINETHRHGCLA----- 106  
Db 63 SSNSLIPNQRGGKTHSDSLANLLOAHRIRDTLDVHLVPVSIFIGRTPDROSGWFAVLSE 122  
Qy 107 -----RRLSVFLFIQERDVH 121  
Db 123 NWALVGRFRLALLAVLNGRNTIVCFAPPISVRQTLNEGLPPERTLKLQRLVLETHFRIR 182  
Qy 122 KGMFATNVT-----ENVLNSRVOEAEIAEVAELNPDGSAQOQSKAVNKKAKRILO 175  
Db 183 ETVIGDLSFRLLDVNLATEAVREAI---AAQAKEDGT-----DLSETWKAQAYAW 233  
Qy 176 EMVATVSPAMIRLTGWLKLFNSFFNQHIGKQLEEMVKAATETNLPFLPLVPHRSHID 235  
Db 234 EIAADYSSPVIRGADFUSFVHMVNRIVAGVLIIH---HVDSPK-ETAPGHEVVYVPSHSHID 290  
Qy 236 YLLLTFTLFCNTHKAPVIAAGNINIPFTLHKLGGFFIRRLDETDPGRKDVLYRAL 295  
Db 291 YMLLSYCLYCGGIVLPHIVAGINLNLPIVGTILKCKGAFFIRSI-----KGNMYSIV 344  
Qy 296 LHGHIVELLRQOQFLEIFLEGTRSRSGKTS CARAGLLSVVVDTLSTNVIPDILIPVGIS 355  
Db 345 LSEVVAQVAGGYSLEYFIEGSRSGTGRLLQPKGMMMTLQAFLRPRPVLFPQIYIG 404  
Qy 356 YDRIIEG-HYNGEOLGPKPKNESLWSVARGVIRMLRKNYGCVRVDPAPQPSLKEYLESQS 414  
Db 405 YEKLIETSYLDELSGSPKESITWRLFNIPKLYKQYQVNVNFGEPITALNDVL----- 460  
Qy 415 QKPVSAALLSLEQALLPAILRSPSDAAGRDTSINESRNATDESRLRRRLIANLAEHILF 474  
Db 461 -----AELAPEWEGQALNENEPALWS-----STVNHARQIQT 494  
Qy 475 TASKSCAIMSTHIVACLLYHRQGDISTLVEDFFVMKEEVLARDPDLGSG----- 527  
Db 495 RINSADVNPINLALLALSTPKHAMGEADLIQITLCKKILL-----ELFYSNRVTVTPH 550  
Qy 528 NSEVDVMAHLOLGNCVTIHTSRNDEFFITPTTVPSPFELNPNYNGVLHVFIMEAIIA 587  
Db 551 TPERITAHAEQI---NILTRVHHPLGDLVLRDGNV-----LLSYFRNNVLHFTASAWVA 604  
Qy 588 CSLVAVLNKRGIGLGPSTPPLNISQEOQVRAKASCLVLLSNEGTSILPCQTFYQVCHETV 647  
Db 605 CCFKN---NRR-----MSRIALIRLGVGMYPFLQAEFLPWTEQFAHQHQOVI 650  
Qy 648 GKFTYQILTVABHDQEDISPSLAEOQWDKLPPELSWSRSDDEDESDFGEEQRDCYLK 707  
Db 651 ELFVREGLLSAGNEE-----DPLTRNTSQTDE----- 679  
Qy 708 VSOSKEHQOQITFLQRLGLPLEA-----YSSAAIFVHNFSGVPPEPYLQKLHKYLITR 762  
Db 680 -----VFLRAISHLOQAERYITISILVKNPGTILSASE-----LESQCLA 724  
Qy 763 TERNVAVYASATVCLVKNVAKMFKDGVFKETKQK 798  
Db 725 AQLSLLYASTA-----PEFFDKLGRFGFIQK 751

## RESULT 8

US-10-369-493-13643  
; Sequence 13643, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 13643  
; LENGTH: 833  
; TYPE: PRT  
; ORGANISM: Pseudomonas fluorescens  
; US-10-369-493-13643

Query Match 11.4%; Score 490; DB 12; Length 833;  
Best Local Similarity 22.4%; Pred. No. 1.5e-39;  
Matches 168; Conservative 149; Mismatches 264; Indels 170; Gaps 22;

Qy 88 LRNVYINETHRHGWLARRLSYVLFIOERDVHKGMPA-----TNVTENVLNSRVOE 141  
Db 178 LRELIHNKGBHER---TVRMAQRILRVHFRNLKAIVGPDISHRRNLVKGLLNQLPLVKQ 233  
Qy 142 AIAEVAELNPDGSAQOQSKAVNKKAKRILOEMVATVSPAMIRLTGWLKLFNSFF 201  
Db 234 AILD-----EAERENISEKAKAALRYGNEIASDYTYTAIR-----FLEVLSWF 279  
Qy 202 WNIQIHKG---OLEMVKAATETNLPFLPVHRSHIDYLLTFLFCNTHKAPVIAAGN 257  
Db 280 WN-KIYDGIKVNHIEGVQKVAQGH-EVIYVPCRSHIDYLLSYLLFRNGLTPPHIAAGI 337  
Qy 258 NLNIPFTSLHKLGGFFIRRLDETDPGRKDVLYRALHGHIVELLRQOQFLEIFLECT 317  
Db 338 NLNMPVIGLSLRRGGAFMRFTKGNP-----LTSVFNEXYHLTLFTKGFVPEYVEGG 391  
Qy 318 RRSRGKTS CARAGLLSVVVDTLSTNVIPDILIPVGISYDRIIEGH-YNGEOLGPKPKNE 376  
Db 392 RRTGEMLPQTKGMLAITLSRLSRMPVIFVPPVIGYVERVLEGTYLGEALGASKKE 451  
Qy 377 SLWSVARGVIRMLRKNYGCVRVDPAPQPSLKEYLESQSOKPVSAALLSLEQALLPAILPS- 435  
Db 452 SIFDIPK-VIGALKQRFQGVAVNFGPEPIKLAFLDSEQ-----PGW 491  
Qy 436 RSDAAGRDTSINESRNATDESRLRRRLIANLAEHILFTASKSCAIMSTHIVACLLYR 495  
Db 492 ROELGPOPKPWLNETTN-----RUGEKVAQHL-----NEAAINPNVLVALLST 539  
Qy 496 HRQGDISTLVEDFFVMKEEVLARDPDLGFS-----GNSDEVVMAH--IQ 538  
Db 540 TRALD-----DRAMARVLDDLALLRKVPYSPHTTLPBGDGRAUHEHVKMD 587  
Qy 539 LLGNVCVTIHTSRNDEFFITPTTVPSPFELNPNYNGVLHVFIMEAIIACSLYAVLNKRG 598  
Db 588 LL-----SEQNDALGKILYLDQNAVLMTYRNNVLHIFALPALLASFFQST----- 634  
Qy 599 LGGPTSTPPLNISQEOQVRAKASCLVLLSNEGTSILPCQTFYQVCHETVGFQYQILTV 658  
Db 635 -----SRMSREQLRYTRALPYLOSELFIWTLDELDAVIDDOWLEAFVEQGLLRF 685  
Qy 659 AEHDDQEDISPSLAEOQWDKLPPELSWSRSDDEDESDFGEEQRDCYLKVSOSKEHQOQFI 718  
Db 686 -----EKDYLAPSPSSRHVLL 703  
Qy 719 TFLQRLGLPLEAYSSAAIFVHNFSGVPPEPYLQKLHKYLITRTE-----RNVAVVAESA 774  
Db 704 TLLSKSIQTLQRFYMTVSVLLNSGQNSISASELEDLCTVMAQRSLHGLNAPPEFDKS 763  
Qy 775 TYCLVKNVAKMFKDGVFKETKQKRVSVLEL 805  
Db 764 ---LFRHF:QTLDDLVLRRDEAGKLSYHEL 791

## RESULT 9

US-10-369-493-9240  
; Sequence 9240, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:

```

651 ELFVREGLLLSAGNEE-----DPLRNTSQUDE----- 679
Db
708 VSQKEHQOFTIFQRLGLPLEA-----YSSAAIFVHNFSGVPPEYLQKLHKYLIR 762
QY
680 -----VFLRAISHLQOAFERYITISLVKNGPGLTASE-----LESICQLA 724
Db
763 TERNVAVAESATCYLVKNAVMPKDIGVFKETKQK 798
QY
725 AQRSLLYASTA-----PEFDFKGLFRGFIQK 751
Db

```

## RESULT 10

```

/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTIVITY
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 16014
/ LENGTH: 772
/ TYPE: PRT
/ ORGANISM: Xanthomonas campestris
/ US-10-369-493-16014

```

Query Match 10.6%; Score 452; DB 12; Length 772;  
Best Local Similarity 25.2%; Pred. No. 9.2e-36;  
Matches 145; Conservative 113; Mismatches 223; Indels 94; Gaps 17

QY	107	RLSLVLFQIBRDVHKGMFATNVT-----ENVLNSRSRVOEATAEVAEAINPDGSAQQOQS	160
Db	186	FKLOVRFLFRRIRREAVIGPDLSTRELLVDVOLAADSVEAATQAKRDN-----S	237
QY	161	KAVNKVKKAKRILIQEWMVATVSPAMILITCGWLLKLFNSPFWMIQIHKGOLEWKAATET	220
Db	238	KPVD-AMRKAHAYAEWAADYSSPVVASGELLTHVWNRIYAGVLVH--HLDKLKOQAPG	294
QY	221	NLPLFLFVHRSHIDYLLLTPIFLFCHNIKAPYIASGNLNIPIFSTLIHLKGLGFFRRRL	280
Db	295	H-EVYVYFESHSHMDYLLLSYLLYGERGIVPPHIVAGINLNPVVGTLRKGGAFFRRSI	353
QY	281	DETDPGRKDVLYRALLHGHIVELLRQOQFLEIFLEGTTRSRSKTSNARAGLLSVVVDTLS	340
Db	354	-----RGNALYSAVLSYEAQVLVAGGYSIEYFVGGRSRTGLLOPKGGMAMTLRAYL	407
QY	341	TNVPIDILIIPVGIYSRIIEGH--YNGEQLGKPKNESLWSVARGVIMLRKNVGCVRVD	399
Db	408	RQRPKPVLFQPVYGYEKLMEGNSYLDLGTGRPKESIWGLLWSYKPVKLQNYGVQVWN	467
QY	400	FAQPFSLKXELYESO-----OKPYSALLSLEQALLPAILPSRPSDAADEGRDTS	448
Db	468	FGEPTALNDVLAKHAPENWDEPLDPDEKPT-----WLAPAV-----DTLATQIQTR	513
QY	449	INESENATDESIRRRRLIANLAEHLFTASKSCAIMSTHIVACLLLYHRHQCIDLSTLVED	508
Db	514	INCAADVNPINLALALLSTPKHAMEAD-----LIAQIELCKKLAEMPYSRDTVT--	566
QY	509	FFVMKEEVLARDPDLGFGSGNSEDVVMHAIQLLGNCVTIHTSRNDEFFITPSTTVPVSPE	568
Db	567	-----PHTPAR-----IITHAEI--NVLTRVSHPLGDVLVSGDTAV-----L	603
QY	569	LNFYSGNGLVHFVIMEAIIACSLAVLNKRGLGGTSTPPNLIISQEOQLVRAAKSLYLSN	628



646 GAAEDRKS FV KVALETG RAEYHAGRIITASESLAKVTILENAVAFLL 690

Db

RESULT 14

US-09-712-363-244

;; Sequence 244, Application US/09712363

;; Patent No. US20020164588A1

;; GENERAL INFORMATION:

;; APPLICANT: Eisenberg, David

;; APPLICANT: Rotstein, Sergio H.

;; APPLICANT: Marcotte, Edward M.

;; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND

;; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS

;; FILE REFERENCE: 07419-032001

;; CURRENT APPLICATION NUMBER: US/09/712,363

;; CURRENT FILING DATE: 2000-11-13

;; PRIOR APPLICATION NUMBER: PCT/US00/02246

;; PRIOR FILING DATE: 2000-01-28

;; PRIOR APPLICATION NUMBER: 60/179,531

;; PRIOR FILING DATE: 2000-02-01

;; PRIOR APPLICATION NUMBER: 60/117,844

;; PRIOR FILING DATE: 1999-01-29

;; PRIOR APPLICATION NUMBER: 60/118,206,

;; PRIOR FILING DATE: 1999-02-01

;; PRIOR APPLICATION NUMBER: 60/126,593

;; PRIOR FILING DATE: 1999-03-26

;; PRIOR APPLICATION NUMBER: 60/134,093

;; PRIOR FILING DATE: 1999-05-14

;; PRIOR APPLICATION NUMBER: 60/134,092

;; PRIOR FILING DATE: 1999-05-14

;; PRIOR APPLICATION NUMBER: 60/165,124

;; PRIOR FILING DATE: 1999-11-12

;; PRIOR APPLICATION NUMBER: 60/165,086

;; PRIOR FILING DATE: 1999-11-12

;; NUMBER OF SEQ ID NOS: 292

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 244

;; LENGTH: 789

;; TYPE: PRT

;; ORGANISM: Mycobacterium tuberculosis

US-09-712-363-244

Query Match 5.7%; Score 245; DB 10; Length 789;

Best Local Similarity 22.0%; Pred. No. 7.4e-15;

Matches 112; Conservative 95; Mismatches 206; Indels 96; Gap

QY 116 QERDV-----HKGMPATNTVENLNSRVQERAEVAAE-----LNPDSGA

Db 149 QWRDTTVAEHRKDRFAQFVSRRAL-----LALARAERYILGPOYKSPRLVLPKPEMLA

QY 159 QSKA-VNKVK---KKAKRILOQMV---ATVSPAMIELTGWVLLKLFNSPFWNIIQIHKI

Db 201 RFRAGLDRIIPATVEDAGKMDLDELTGWSQVSDVLVSLGRLASRGDDEF---DYDEE

QY 211 LEMVKAATETNLPLFLPVRHSHDYLLLLFILFCHNIKAPYIASGNLNIPIFTSLII

Db 258 VAAMEAALAEH-PAVLFSHRSYIDGVVVPVAMQDNELPVPVMEGGINLSFGLMGLPLM

QY 271 LGGFTPIRRRLDETPDGRKDVLYRALLHGHIVELLRQQOFLFILEGTRSRSGKTSKAR

Db 317 SGMIPIRRNIGNDP-----LYKVLVEYGVVVEKRFNLSWSIEGTRSRRTGKMLPKP

QY 331 LLSVVVDTLSTNVIPDILIIIPVGISYDRIITE-CHYNGEQIGKPKKNESLWSVARGVIR

Db 371 LMSYVADAYLQGRSDDLILQGVSVICFDOLHEITEYAAVARGAEKTPGRLRWLYNFKA

QY 390 RKNVGCVRVDFAQPFSLKEYLESQKVPVSALLSLEQALLPAILPSRPSDAADEGRDT

Db 431 ENFCKIYVRPEAVSMRQYL-----GAPHELTV

QY 450 NESRNATDESIRRLRIANLAEHILFTASKCAIMSTHIVACLILLYRHROGIDISTLVE

Db 461 -----DPAKRIALQKMSFEVAWRILQATPVTATGLVSAALLTTTGTATLTDOLHIA

QY 508 ----DFPMKEVLARDFDLGFGSNDSDVVMHAIQLL--GNCVTIHTSRNDEFFITPST 561  
Db 514 QDSLDYLERKQSPVST--ALRLSRGVRRAADALSNGHPVTRVDSGREPVWYIAPDD 570  
QY 562 TVPSVFELNFYSNGVLHVFMFIMEAIIACSL 590  
Db 571 EHA AAA----FYRNSVIAHFLTSETSELAL 595  
RESULT 15  
US-10-094-749-1905  
; Sequence 1905, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOKYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1905  
; LENGTH: 1095  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-094-749-1905

Query Match 3.0%; Score 128; DB 12; Length 1095;  
Best Local Similarity 19.2%; Pred.No.0.0083;  
Matches 155; Conservative 118; Mismatches 329; Indels 204; Gaps 33;  
QY 114 FIOERDVHKGMAFNVTENVLNSRVOEAI AEVAAELNPDGSAQQOQSAKVNKVKKAKRI 173  
Db 349 FLQNLVPHRSVSVTILEVKNVSHVDHVTQGLVEL--GFILMDSYG-----PKKV 398  
QY 174 LOEWVATVSPAMRLTGWLLK-----LFNSFFWNIQIHKGQLEMV--KAATETNLPFLF 226  
Db 399 LDGKTETSPSLRMPNQHACKLGANILLETFTKIHMIROEILEQVLNRRVVTASSPI-- 456  
QY 227 LPVHRSHIDVLLTFFILFCHNIKAPYIASGNLNI-----PI 263  
Db 457 -----SHFDLLSNVIMY-----APLVQSCSKVTEAFDYLSPLOTVORLLKAVQPL 506  
QY 264 FST-----LIHKLGGFFIRRLDETPDGRKDVLYRALHGHIVELLRQOQFLEIFLEG 316  
Db 507 LKVSMSMRDCLILVLRKAMFANQL-----DARKSAVAGFLL-----LLKNFKVL-----G 551  
QY 317 TRSRSGKTS CARGLLSVVVDTLSTNVIPDILIPVGISYDR-----IIEGHYN 365  
Db 552 SLSSQCSQSLSVQVHVHSHYNSVANETFTCLEIMDSLRCLSQOQADVRLMLYEGFYD 611

QY 366 GBQLGKPKKNESLW--SVARGVIRMLRKNYGCVRVDPAPQFSLKEYLESQS-----QKPVS 419  
Db 612 -----VLRENSQLANSVMQTLTLLSQLKQFYE--PKPDLPLPELKEACTLTQGDKISLQEPD 665  
QY 420 ALLSLEQALLP-----AILPSRPSDAADGRDTSINESRNATDESRLRLIANLAELHIFT 475  
Db 666 YLLCCIOHCLAWYKNTVIPLOQGEHEE--EEBAFYEDLDLLESITNRMIKSELED--FE 722  
QY 476 ASKSCAINTSHIVACLLLYRHRQGDLSLVEDDFVMK--BEVLARDFDL--GFSGNSEDV 532  
Db 723 LDKSAD-----FSQSTSIGIKNNISAFVLMGFCEVLIIEYNSISISSFKNRFD 770  
QY 533 VMHAIQLLGNVCVTIHTSRNDEFFITPSTVPSVFMFELNFYSNGVLHVFMFIMEAIIACSLYA 592  
Db 771 ILSLFWCYKKLSLILNEKAGAKTKWANKTSDLSLSMKFVSS-----LLTALFR 819  
QY 593 VLNKRGLGGPTSTPPNLISQEQOLVRKAASLCVLLSNEG--TISLPCQTFYQVCHETVGKF 650  
Db 820 VLLWRYTSIPTSV-----EESGKKEKGKSISSLCLLEGLOKIFSAVQOQFYQ---PKIQOF 870  
QY 651 IQYGILTVAEHDDQEDISPSLAE-----QOWDKKLPEPLSWRSDEDESDSDFGEEQDC 704  
Db 871 LRALDVTKEGEERDADSVTQRTAFQIRQORSLLNLS--SQEEDFNS-----KEA 922  
QY 705 YLKVS-----OSKEHQQFITTFLORLLGLPLLEYSSA 735  
Db 923 LLLVTVLTSLSKLLBPSSPQFVQMLSWTSKICKENSREDALFCKSLMNLFLSLHVSYSKP 982  
QY 736 AIF-----VINFGPVPEPEYLOKLHKYLTIRTRNVAVVYAESATYCLVKNVAKMPKD 788  
Db 983 VILLRDLSDQIHGHLGDIDQDVEVEKTNHFAIV-----NLRTAAPTVCLLVLVLSQAEKVLEE 1038  
QY 789 IG-----VFKETKQKRVSULEL 805  
Db 1039 VDMLITKLKGQVSQETLSVSPGVSEL 1064

Search completed: January 10, 2004, 20:29:32  
Job time : 62 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2004, 19:36:42 ; Search time 70 Seconds  
(without alignments)  
3052.394 Million cell updates/sec

Title: US-09-935-290-2  
Perfect score: 4280  
Sequence: 1 MDESALTGTIDVSYLPHSS.....FLPQCNRQKLEYILSFVVL 828

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL 23.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phage.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriap.\*
  - 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	4265	99.6	828 4 Q8N1G6	Q8N1G6 homo sapien
2	3993.5	93.3	827 11 Q8VCT2	Q8VCT2 mus musculu
3	834	19.5	850 5 Q9Y137	Q9Y137 drosophila
4	831.5	19.4	757 5 Q8IMM7	Q8IMM7 drosophila
5	831.5	19.4	786 5 Q8IMM8	Q8IMM8 drosophila
6	442	10.3	809 16 Q8E809	Q8E809 shewanella
7	424	9.9	678 11 Q8R390	Q8R390 mus musculu
8	399.5	9.3	809 16 Q8DD48	Q8DD48 vibrio vuln
9	393	9.2	680 4 Q8BWC2	Q8BWC2 homo sapien
10	361	8.4	794 16 Q8D373	Q8D373 wiggleswort
11	353.5	8.3	663 5 Q9NHF2	Q9NHF2 drosophila
12	349.5	8.2	724 5 Q9VBQ6	Q9VBQ6 drosophila
13	337	7.9	671 5 Q9U2E8	Q9U2E8 caenorhabdi
14	335.5	7.8	320 11 Q8BRZ9	Q8BRZ9 mus musculu
15	178	4.2	809 16 Q8EYF2	Q8EYF2 leptospira
16	150	3.5	287 16 Q8YMR2	Q8YMR2 anabaena sp

17	140	3.3	3144	4 Q9UQB7	Q9UQB7 homo sapien
18	138.5	3.2	249	16 Q8PF86	Q8PF86 xanthomonas
19	135.5	3.2	239	16 Q97L74	Q97L74 clostridium
20	134	3.1	240	16 P73054	P73054 synechocyst
21	130.5	3.0	249	2 Q8EY25	Q8EY25 xanthomonas
22	129.5	3.0	249	16 Q8P3Q9	Q8P3Q9 xanthomonas
23	129.5	3.0	1084	4 Q9NV11	Q9NV11 homo sapien
24	128	3.0	356	10 Q8GXU8	Q8GXU8 arabidopsis
25	122	2.9	654	4 Q86IV0	Q86IV0 homo sapien
26	120	2.8	1129	3 Q8J1G4	Q8J1G4 ashbya goss
27	118.5	2.8	906	4 Q86ST0	Q86ST0 homo sapien
28	118	2.8	651	11 Q8K113	Q8K113 mus musculu
29	118	2.8	1258	4 Q00539	Q00539 homo sapien
30	118	2.8	1258	11 Q08982	Q08982 mus musculu
31	116	2.7	1262	3 Q13327	Q13327 candida alb
32	115.5	2.7	564	10 Q8L4Z8	Q8L4Z8 oryza sativ
33	115	2.7	651	11 Q9J178	Q9J178 mus musculu
34	115	2.7	4133	13 Q9DE12	Q9DE12 gallus gall
35	115	2.7	4134	13 Q8QXK4	Q8QXK4 gallus gall
36	114	2.7	675	17 Q8TNP4	Q8TNP4 methanosarc
37	113	2.6	1231	4 Q8N3U4	Q8N3U4 homo sapien
38	112.5	2.6	1885	2 Q8RJY4	Q8RJY4 stigmatella
39	112	2.6	344	10 Q8W3E1	Q8W3E1 oryza sativ
40	112	2.6	822	5 P90531	P90531 dictyosteli
41	112	2.6	2273	3 Q42823	Q42823 saccharomyc
42	111.5	2.6	635	10 Q8S0J1	Q8S0J1 oryza sativ
43	111.5	2.6	805	5 Q95RK7	Q95RK7 drosophila
44	111.5	2.6	921	5 Q9VJ29	Q9VJ29 drosophila
45	111	2.6	292	16 Q9KIH9	Q9KIH9 rhizobium m

ALIGNMENTS

RESULT 1

Q8N1G6 PRELIMINARY; PRT; 828 AA.

ID Q8N1G6; AC Q8N1G6; DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Hypothetical protein KIAA1560.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Strausberg R.;

RL Submitted (MAY-2002) to the EMBL/GenBank/DBDJ databases.

DR EMBL; BC030783; AAH30783.1; --

DR InterPro; IPR002123; Acyltransferase.

DR SMART; SM00563; PLG; 1.

KW Hypothetical protein.

SQ SEQUENCE 828 AA; 93874 MW; 55F6CF7EBCFE227A CRC64;

Query Match	99.6%	Score 4265;	DB 4;	Length 828;
Best Local Similarity	99.6%;	Pred. No. 0;		
Matches 825;	Conservative 1;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1	MDESALTGTIDVSYLPHSSYSVGRCKHTSEEWGECGFRPTVFRSATLKWKESLMSRRK	60	
Db	1	MDEALTGTIDVSYLPHSSYSVGRCKHTSEEWGECGFRPTVFRSATLKWKESLMSRRK	60	
Qy	61	PFVGRCCVSCPTQSWDKFFNPSIPSLGLRNVIYINETHTRHGMRLARLSYVLFQERDV	120	
Db	61	PFVGRCCVSCPTQSWDKFFNPSIPSLGLRNVIYINETHTRHGMRLARLSYVLFQERDV	120	
Qy	121	HKGFMATNTENVLNSRVRQAIAEVLAELNPDGSAQQOSKAVNKVKKKAKRILOEMVAT	180	
Db	121	HKGFMATNTENVLNSRVRQAIAEVLAELNPDGSAQQOSKAVNKVKKKAKRILOEMVAT	180	



Qy	181	VSPAMIRLTGWLLKLFNSFFWNIQIHKGQLEWMAKATETNLPFLFVHRSHIDYLLLT	240
Db	181	VSPAMIRLTGWLLKLFNSFFWNIQIHKGQLEWMAKATETNLPFLFVHRSHIDYLLLT	240
Qy	241	FILFCHNIKAPYIASGNLNIPISTLIHKLGGFFIRRRRLDETPDGRKDVLYRALLHGI	300
Db	241	FILFCHNIKAPYIASGNLNIPISTLIHKLGGFFIRRRRLDETPDGRKDVLYRALLHGI	300
Qy	301	VELLRQOQFLEIFLEGTRSRGKTSKARAGLLSVVVDLTSTNVPDILIIIPVGISYDRII	360
Db	301	VELLRQOQFLEIFLEGTRSRGKTSKARAGLLSVVVDLTSTNVPDILIIIPVGISYDRII	360
Qy	361	EGHYNGQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPPSKLEYLESQKPVSA	420
Db	361	EGHYNGQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPPSKLEYLESQKPVSA	420
Qy	421	LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC	480
Db	421	LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC	480
Qy	481	AIMSTHIVACLLYRHRQIDLSLTVDFVFMKEEVLARDFDLGFSGNSDVVMAIQLL	540
Db	481	AIMSTHIVACLLYRHRQIDLSLTVDFVFMKEEVLARDFDLGFSGNSDVVMAIQLL	540
Qy	541	GNCVTIHTSRNDEFFITPTSTVPSVFNLFYNGVLHVFMIAIACSIVAVLNKRGGLG	600
Db	541	GNCVTIHTSRNDEFFITPTSTVPSVFNLFYNGVLHVFMIAIACSIVAVLNKRGGLG	600
Qy	601	GPTSTPNLISQQLVKAASLCYLLSNEGTSISLPQCTFYQVCHETVKGFIQYGLTVAE	660
Db	601	GPTSTPNLISQQLVKAASLCYLLSNEGTSISLPQCTFYQVCHETVKGFIQYGLTVAE	660
Qy	661	HDDQEDISPSLAEOQWKKLPEPLSWRSDEDEDSDFGEORDCYLKVSQSKHQOIFITF	720
Db	661	HDDQEDISPSLAEOQWKKLPEPLSWRSDEDEDSDFGEORDCYLKVSQSKHQOIFITF	720
Qy	721	LQRLGLPALLEAYSSAAIFVHNFSGVPPEPEYLOKLHKYLIITRERNVAVVAESATYCLVK	780
Db	721	LQRLGLPALLEAYSSAAIFVHNFSGVPPEPEYLOKLHKYLIITRERNVAVVAESATYCLVK	780
Qy	781	NAVFMFKDIGVFKETQKRVSVLELSSTFLPQCNRQKLLLEYILSFVVL 828	
Db	781	NAVFMFKDIGVFKETQKRVSVLELSSTFLPQCNRQKLLLEYILSFVVL 828	

RESULT 2

Q8VCT2	ID	Q8VCT2	PRELIMINARY;	PRT;	827 AA.
AC	Q8VCT2;				
DT	01-MAR-2002	(TrEMBLrel. 20, Created)			
DT	01-MAR-2002	(TrEMBLrel. 20, Last sequence update)			
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)			
DE	Glycerol-3-phosphate acyltransferase, mitochondrial.				
GN	GPAM.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Kidney;				
RA	Strauberg R.;				
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC019201; AAH19201.1; -.				
DR	MGD; MGI:109162; Gpam.				
DR	InterPro; IPR002123; Acyltransferase.				
DR	SMART; SM00563; Plsc; 1.				
KW	Transferase; Acyltransferase.				
SQ	SEQUENCE 827 AA; 93704 MW; 4C177AA15374EE9B CRC64;				

Query Match 93.3%; Score 3993.5; DB 11; Length 827;  
 Best Local Similarity 92.8%; Pred. No. 9e-317;  
 Matches 768; Conservative 30; Mismatches 29; Indels 1; Gaps 1;

Qy	1	MDBSALTGTIDVSYLPHSSEYSGRCKHTSBWEGCGFPTVFRSATLKWESLSMRKR	60
Db	1	MEESSVTGTIDVSYLPHSSEYSLGRCKHTSEDWDCGKPTFFRSATLKWESLSMRKR	60
Qy	61	PVGRCCYCTQSDWKFFNPSPISGLRNVIYINETHTRHGWLARRLSYVLFIOERDV	120
Db	61	PVGRCCYCTQSDWERFFNPSPISGLRNVIYINETHTRHGWLARRLSYVLFIOERDV	120
Qy	121	HGMPATNTENVLNSRRVQEAIAEVAELNPDGSAQOQSKAVNKVKKAKRILQEWAT	180
Db	121	HGMPATSVTENLSSSRVQEAIAEVAELNPDGSAQOQSKAIQVKRKARKILOEWAT	180
Qy	181	VSPAMIRLTGWLLKLFNSFFWNIQIHKGQLEWMAKATETNLPFLFVHRSHIDYLLLT	240
Db	181	VSPAMIRLTGWLLKLFNSFFWNIQIHKGQLEWMAKATETNLPFLFVHRSHIDYLLLT	240
Qy	241	FILFCHNIKAPYIASGNLNIPISTLIHKLGGFFIRRRRLDETPDGRKDVLYRALLHGI	300
Db	241	FILFCHNIKAPYIASGNLNIPISTLIHKLGGFFIRRRRLDETPDGRKDVLYRALLHGI	300
Qy	301	VELLRQOQFLEIFLEGTRSRGKTSKARAGLLSVVVDLTSTNVPDILIIIPVGISYDRII	360
Db	301	VELLRQOQFLEIFLEGTRSRGKTSKARAGLLSVVVDLTSTNVPDILIIIPVGISYDRII	360
Qy	361	EGHYNGQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPPSKLEYLESQKPVSA	420
Db	361	EGHYNGQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPPSKLEYLESQKPVSA	420
Qy	421	LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC	480
Db	421	PLSLEQALLPAILPSRPNVADHQDLSNESRNPADEAFRRRLIANLAHILFTASKSC	480
Qy	481	AIMSTHIVACLLYRHRQIDLSLTVDFVFMKEEVLARDFDLGFSGNSDVVMAIQLL	540
Db	481	AIMSTHIVACLLYRHRQIHLSLTVDFVFMKEEVLARDFDLGFSGNSDVVMAIQLL	540
Qy	541	GNCVTIHTSRNDEFFITPTSTVPSVFNLFYNGVLHVFMIAIACSIVAVLNKRGGLG	600
Db	541	GNCVTIHTSRNDEFFITPTSTVPSVFNLFYNGVLHVFMIAIACSIVAVLNKRCSG	600
Qy	601	GPTSTPNLISQQLVKAASLCYLLSNEGTSISLPQCTFYQVCHETVKGFIQYGLTVAE	660
Db	601	GSAGGLGNLISQQLVKAASLCYLLSNEGTSISLPQCTFYQVCHETVKGFIQYGLTVAE	660
Qy	661	HDDQEDISPSLAEOQWKKLPEPLSWRSDEDEDSDFGEORDCYLKVSQSKHQOIFITF	720
Db	661	QDDQEDVSPGLABQWQWKKLPE-LNWRSDDEDESDGEEQDCYLKVSQSKHQOIFITF	719
Qy	721	LQRLGLPALLEAYSSAAIFVHNFSGVPPEPEYLOKLHKYLIITRERNVAVVAESATYCLVK	780
Db	720	LQRLGLPALLEAYSSAAIFVHNFSGVPPESEYLOKLHKYLIITRERNVAVVAESATYCLVK	779
Qy	781	NAVFMFKDIGVFKETQKRVSVLELSSTFLPQCNRQKLLLEYILSFVVL 828	
Db	780	NAVFMFKDIGVFKETQKRVSVLELSSTFLPQCNRQKLLLEYILSFVVL 827	

RESULT 3

Q9V137	ID	Q9V137	PRELIMINARY;	PRT;	850 AA.
AC	Q9V137;				
DT	01-NOV-1999	(TrEMBLrel. 12, Created)			
DT	01-NOV-1999	(TrEMBLrel. 12, Last sequence update)			
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)			
DE	BCDNA:GH07066 protein.				
GN	BCDNA:GH07066 OR CG5508.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				





RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazeg R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Brill J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Gladok A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,  
RA Hostin N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin G., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhou L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RA "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Celinker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreria S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
RT "Sequencing of *Drosophila melanogaster* genome";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celinker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
RT "Annotation of *Drosophila melanogaster* genome";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003763; AAN14132.1;  
SQ SEQUENCE 786 AA; 88312 MW; 6A3084B77993620 CRC64;

Query Match 19.4%; Score 831.5; DB 5; Length 786;  
Best Local Similarity 28.9%; Pred. No. 9.1e-59;  
Matches 215; Conservative 165; Mismatches 262; Indels 101; Gaps 19;  
QY 129 VTEVNLSSRVOEIAEVAAB-----LMPDGAQQQSK----- 161  
DB 69 VTEIHNDAVLKHAQAAEQTLREORYAKNGHRLTRTSASGDQEAQEEDEKRGISYQA 128  
QY 162 AVNKKVKKAKKILEMVAATVSPAMIRLTGWLLKLFNSFFWNIQIHKQLEWVKAATSTN 221  
DB 129 ILRQKQRAISILKDMGSTLNGGLATSTWLYKLLPCFLSGVVTNTKQIEMLKATATERS 188  
QY 222 --LPLELPVHRSHIDYLLTFLIFCHNIKAPYIASGNNLNIPIFSLIHLKGFFFRRR 279  
DB 189 PGTELIIFVPLHRSHLDYIMVTWILNTNDRSPLVAAGNNLQIPVFGLLRGLGAFFIKRK 248  
QY 280 LDETPDGRKQVLYRALLGHIVELLRQOQFLEIFLEGTRSRSGKTSCARALLSVVDTL 339  
DB 249 IDPV-EGKQDVLRYAALHLVLTALKQCHNVFFIEGGRTRTKGPKMPKGGILSVINAF 307  
QY 340 STNVIPDILLIPVGISYDRIIEGHYNGEOLGPKKNESLMSVARGVIRMLRKNYGCVRVD 399  
DB 308 MDGSIPLLIPVSVNYERLVDFGNVREQGEKKIPESFGKAIISGIWALKASNYGLMKRID 367  
QY 400 PAQPFSLKYLESQSQ-----KPVSAALLSLEQALLPAILSPSPSDAADEGRDT 447  
DB 368 FNEPYSIRELVNSYNKIAREDGNTIAKVYKPSARVLQHNQ-----STSSLYGTDV 416  
QY 448 SINESRNATDESRRRIIANAEHLITASKSCAIMSTHIVACLLLYVHRGID---LST 504  
DB 417 VCEEHRN-----LIESISQVDFCAATSVMTNALAFLLTLFRNGABEQILSE 467  
QY 505 LVEDFFVKSEVLARDFDLGSNGSDVMHAIOLLGNCVTITHTSRNDEFFITPSTVP 564  
DB 468 ALDDL-----RNSLSCGCKDIGFSGSQIVAYACDLGSLVTRSDENGRULIVKAVNSVE 523  
QY 565 SVFELNFYSGVLHVFIMEAIIACSLYAVL-----NKRGLGGPTSTPPNLISEQQLVRKAA 620  
DB 524 SFIELAYYSNMLTPTFPALSSILLTTFHSLLPETENKEAA-----VSRKLLIDTAL 574  
QY 621 SLCYLLNNEGITISLPCTQTFVQVCHETVGKFIQYG-ILTVAEHDDQEDISPISLAEQQMDKK 679  
DB 575 ENCOIYRYEFTLNKPTQVLENLLYQQLDDLLISGCVLTKLHDD---LPNGAE---GRR 627  
QY 680 LPEPLSWRSEED-EDSDFGEEQDRCVLYKSQSEHQOQFITFLQRLGLPLEAYSSAA-- 736  
DB 628 LANYLAECLEDDGYEDYRDGEADEPKLLFASETPSQORYIC---EVLAPFAWTYTVTAQS 684  
QY 737 -IFVHNFSGPVPEYELQKLHKYLIITERNVAVVAESATYCLVKNAVMPKDIQVFKET 795  
DB 685 LQILHKNS--MLESEFISFVINDLSKVKRGSCIAESISTDSVRNCLLLEKKSVEVC 742  
QY 796 KQKRVSVLEISSTFLPQCNRQKL 818  
DB 743 NQCGMRLISLNTLY--EMSRESL 763  
RESULT 6  
Q8E8Q9  
ID Q8E8Q9 PRELIMINARY; PRT; 809 AA.  
AC Q8E8Q9;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
GN Glycyl-3-phosphate acyltransferase.  
OS Shewanella oneidensis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
OC Alteromonadaceae; Shewanella.  
OX NCBI\_TaxID=70863;  
RN [1]  
RP SEQUENCE FROM N.A.



```

QY 766 NVAVVAESATYCLVKNVAKMFKDIGVFKETK 796
DB 613 GSSQCYDALSSSELQNALAAAFVRLGVVEKK 643

RESULT 8
Q8DD48
ID Q8DD48 PRELIMINARY; PRT; 809 AA.
AC Q8DD48;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glyceral-3-phosphate O-acyltransferase.
GN VV11165.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016800; AAC0936.1; -.
KW Acyltransferase; Transferase; Complete proteome.
SQ SEQUENCE 809 AA; 90596 MW; OE04PFOB4980DCDC CRC64;

Query Match 9.3%; Score 399.5; DB 16; Length 809;
Best Local Similarity 22.4%; Pred. No. 1.7e-23;
Matches 180; Conservative 145; Mismatches 223; Indels 257; Gaps 38;

QY 84 PSLGLRNVIYNEITH-----TEHGWLARRLSYLVFIQERDVHKG-----NPAIVNT 130
DB 169 PVWSLR-----YQAOSHGTSSIAHK---LARVARIHFSRQKLAAGDPDLPSSQVLF 218
QY 131 ENVLNRSRVQBAIAEVAELNPDGSAQQOSKAVNKKAKRILOEMVATVSPAMI----- 186
DB 219 -RLMKSPIAEQAIIE-----EAKKNISMEEKARQADIMEIAADFSYLVKQGD 268
QY 187 RLGTGWLKLFNSFFWNTQIHKGQLEMKVAAETNLT-----PLFLPVHRSHIDYLLTF 241
DB 269 RLLGLW-----WN-KLYQG-LNINNAATVRLAQDGHGHEIVVPCVHRSHMDYLLSY 317
QY 242 ILFCHNIKAPIASGNLNI-----PIFSTLHKLGFPFIRRLDETPDGRKDVLYRALLH 297
DB 318 VLYHEGMVPPHIAAGINLNFPPAGPIF---RRGGAFFIRRSF-----KGNRLYSTIFR 367
QY 298 GHIVELLRQQOFLEIFLEGTSRSCKTSCARAGLLSVVVDLTSTNVIPIIDILIPVGISYD 357
DB 368 EYLAEFLPAKYSVEVFSGGSRGRTGLLPAGTKMLAMTIQAWLGNRPVTLVPVYIGYE 427
QY 358 RIIE-GHYNGEQLGPKPKNESLWSVARGVIRMLRK--NYGCVRVDFAPFSLKEYLESQS 414
DB 428 HVMVAVATYAKELGRKKEKEN---AGLVRLTKLRNFGLYNGFYNGEPIPLNQYLYNEHA 483
QY 415 QK-----PVSALLSLEQALLPAILPSPSDAABEGRTSINESRNATDESRLRLIANL 468
DB 484 PEWTKNDIDPMGA-----SRPQ-----WINP-----VVNQL 508
QY 469 AEHILFTASKCAIMSTHIVACLLLYRHRQIDILSTLYED-----PFVMEKEVLARDPDLG 524
DB 509 ANKMWTHINDAANAULTCATTALLASQKALSQKSLHIOECYQLLKNVPYKTYVP 568
QY 525 FSGNSEVYVMAIQI-----LGNCVTITHSRNDEFITPTSTVPVSFELNFSNG 575
DB 569 -SESAAEALVEHAIISLDKFIETDWTGDIISL---DRNQSIIMT-----YRNN 612
QY 576 VLVHFMIEAITACSLYAVLNKRLGGTPTSPNLIISQOLVKAASLCYLLSNEGTTISLP 635
DB 613 IHLFALPSLIA-----QMIIRQEN----- 632

```

RESULT 9  
 Q9BWC2  
 ID Q9BWC2 PRELIMINARY; PRT; 680 AA.  
 AC Q9BWC2;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Glyceral-3-phosphate O-acyltransferase.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA Strausberg R.;  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC000450; AAH00450.1; -.  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 DR SMART; SM00563; Plac; 1.  
 KW Acyltransferase; Transferase.  
 SQ SEQUENCE 680 AA; 77161 MW; A6EC9567D5693476 CRC64;

Query Match 9.2%; Score 393; DB 4; Length 680;  
 Best Local Similarity 25.6%; Pred. No. 4.4e-23;  
 Matches 141; Conservative 113; Mismatches 193; Indels 104; Gaps 24;

```

QY 68 YCTPQSWDKFPNPSIPSLGLRNVIYINETHRHWLARRLSYLVFIQERDVHKGMP-- 125
DB 25 YSKELKKWDEF-----EDIL---BERRH---VSDLKFMKCYTPLYKGTIPC 66
QY 126 -ATNVTENVLNRSRVQBAIAEVAELNPDGSAQQOSKAVNKKAKRILOEMVATVSPA 184
DB 67 KPDIKCSVLNSEEIHVIKQLSKE-----SLQSDVLDREEVSEILDEMSHKLRLG 117
QY 185 MIRLTGWLLKLFNSFFWNTQIHKGQLEMKVAAETNLTPLFLPVHRSHIDYLLTF 244
DB 118 AIRFCATLTKVFKQIFSKVCVNEEGIKQLQALQEH-PVLLPSHRSYIDFLMLSLFLY 176
QY 245 CHNIKAPIASGNL--LNIPFSTLHKLGFFIRRLDETPDGRKDVLYRALLHGHIVEL 303
DB 177 NYDLPVPVIAAGDFLGMKVMGELLRNMGGAFFMER-----TFGNGK--LYWAVFSEYVKTM 230
QY 304 LRQOQF-LEIFLEGTSRSCKTSCARAGLLSVVVDLTSTNVIPIIDILIPVGISYDRIEG 362
DB 231 LRNGAYAEVEFLEGTSRSCKTSLTPKFGLLNIVMEPFKREVFDTYLVPIISYDKILEE 290
QY 363 H-YNGEQLGPKPKNESLWSVARGVIRMLRNKYNVCVRVDEAQPFSLKEYLESQSOKPVSA 421
DB 291 TLVYVELLGVPKPKESTTGLIK-ARKILSNFSGSIHVYFGDPSVLSRLAAGMRMSR----- 345
QY 422 LSLEQALLPAILPSPSDAABEGRTSINESRNA--TDESILRRRL--IANLAHILFTAS 477
DB 346 ---SYNLVPRYIPQKQS-----EDMHAFTVEVAYKMLLOIENM----- 381

```





Db 389 -----ATGRLOFTPMAKQLAEQIDVRLKAHALCPOTMAVAVPMLALQLYI 435  
QY 641 QVCHETVGFQYIGILTVAEHDDQEDISPSLAEQQWDKPLPEPLSWRSDEEDEDSPGEE 700  
Db 436 NCFMFWLARPAYALLAALKEQKQKQSTDIS---YDASL---CALHAHVTTMDALFQHE 489  
QY 701 QRDYLVKVSQSKHEHQOFTIFLQRLGIP-LLEAYSSAAIFVHN----- 741  
Db 490 -----FIIESNREAAEFETHLQLLDERVVEVETSGRINVDNECSHVILAALAPFLCLY 544  
QY 742 -----FSGP---VPPEYLOKL-----HKYLITRTRNVAVVAE 772  
Db 545 YQLVTLRKIPLELEFSNKELLVRVQOHVEQLQOPGASASHVHPYCLALDNLNIAIYAL 604  
QY 773 SATYICLVKNAVMFKDVGFKETKQKRVSVLELSSTFLPQCNQKLEY 821  
Db 605 IQRGYLVKS-----RDSGQMKIATPGKCLRELETQLLEYCQLMPPAQY 648

RESULT 12  
Q9VBQ6 PRELIMINARY; PRT; 724 AA.  
AC Q9VBQ6;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Glyceral-3-phosphate-acyltransferase protein.  
GN DHAP-AT OR CG4625.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=107311132;  
RA Adams M.D., Celiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Fandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Galbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Mohtrefi A.,  
RA Mount S.M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,

RESULT 13

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195 (2000).  
DR EMBL: AE003752; AAF56473.2; --  
DR FlyBase: FBgn0040212; Dhap-at.  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
DR SMART: SMO0563; Pfam; 1.  
KW Acyltransferase; Transferase.  
SQ SEQUENCE 724 AA; 82524 MW; 5972629749CED804 CRC64;  
Query Match 8.2%; Score 349.5; DB 5; Length 724;  
Best Local Similarity 20.8%; Pred. No. 1.7e-19;  
Matches 169; Conservative 135; Mismatches 279; Indels 228; Gaps 31;  
QY 117 ERDVHKGFMATNTVENVL-----SSRQEAETAAEVAE---LNP----- 152  
Db 21 EQEQKTPSAAYWRNFKNITAPGNEASMTREFNPQVAIEFEKYNLPKQLKQHLVRSKL 80  
QY 153 ----DGSAAQQOSKAVNKVKKAKRILOEMVATVSPAMIRLTGWVLL----KLFNSFFWNI 204  
Db 81 RSILEHYAKESGTPKQMERQARALIDIGLDNRWAIIRWCGIAITAIGKRICDGFY--- 137  
QY 205 QIHKGQLEMVKAATETNL-----PLLFLPVHRSHIDYLLTLFTLCHNTKAPYASGN 257  
Db 138 -----VNSASMANVRKMGKCPVLYLPSHRSYMDFILMSYICYVYDIEIPGIAAGM 188  
QY 258 NLNIPF--PSTLIHLKLGFFIRRLDETPDGRKDVLYRALLHGHIVELLRQQOF-LEIFLE 315  
Db 189 DFHSMFGMTWLRKTGAPFMERSFS-----NDELYMDIFREYMYALVANYHIGVEFFIE 242  
QY 316 GTRSRSGKTSARAGLLSVVVDTLSTNVIPIILLIPVGISYDRIIEGH-YNGEQLGKPKK 374  
Db 243 GTRSRNFALVPKIGLLSMALLPYFTGEVDPVNIIVPVSVAYERVLEEQLFVYELGVPKP 302  
QY 375 NESLWSVARGVIRMLR---KNYGCVRVDFAQFSLKEYLESQSKPKVSALLSLEQALLPA 431  
Db 303 KES-----TKGFFKALKIIDERFGKMFIDFGEPISVKEFF----- 337  
QY 432 ILPSRPSDAADGSDTSINESRNATDSLR-----RLIANLAELHILFTASKSCAIST 485  
Db 338 -----GHDSAQRMQRAGVGGHLOKLNQREVELVKQLANEIYIQQORRIVISTF 385  
QY 486 HIVACLLYRHROGIDISTLSTLVEFVFMKEEVLARDFDLFGSGNSEDVVMH---AIQLLGN 542  
Db 386 NLLS--LYASQLAQRSVLDE-----LARG-----VHLKRIFEQLG- 422  
QY 543 CVTITHSRNDEFFITPTSTVPSVFELNFGVGLHVFIEMAIACSLYAVLNKRLGGP 602  
Db 423 ----AHVS-----TTPSSIKADVDAVEIHSNHLF-----ATGRL 454  
QY 603 TSPT---PNLISQOLVR-KAASLCYLLSNEGTTSLPCQTFYQVCHETVGFQYIGILTUV 658  
Db 455 QFTPMARQKLAQEQIDVKRLKAHALCPOTMAVAVPMLALQLYINPCMFELARPAVLLAAL 514  
QY 659 AEHDDQEDISPSLAEQQWDKPLPEPLSWRSDEEDEDSDGEEQRCYLVKVSQSKHEHQOFI 718  
Db 515 KEQNNQKQSTDIS---YDASL---CALHAHVTTMDALFQHE-----FIIESNREAAE 563  
QY 719 TFLQRLGIP-LLEAYSSAAIFVHN-----HKYLITRTRNVAVVAEATYCLVKNVAKMFKDIG 744  
Db 564 THLQQLLDERVVEVETSGRINVDNECSHVILAALAPFLCLYQVLVTLRKIPLELEFSN 623  
QY 745 P---VPPEYLOKL-----HKYLITRTRNVAVVAEATYCLVKNVAKMFKDIG 790  
Db 624 KELLVRVQOHVEQLQOPGASASHVHPYCLALDNLNIAIYALIQRGYLVKS-----RDSG 678  
QY 791 VFKEITKQKVSVLELSSTFLPQCNQKLEY 821  
Db 679 QMKIATPGKCLRELETQLLEYCQLMPPAQY 709



AC	QBEYF2;
DT	01-MAR-2003 (TrEMBLrel. 23, Created)
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Glycerol-3-phosphate dehydrogenase (EC 1.1.1.94).
GN	GPSA3 OR LA4264
OS	Leptospira interrogans.
OC	Bacteria; Spirochaetes; Spirochaetales; Leptospiaceae; Leptospira.
OX	NCBT_taxid=173;
RN	[1]
RC	SEQUENCE FROM N.A.
RP	STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA	Ren S.;
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AE011579; AA051462.1; -
KW	Oxidoreductase; Complete proteome.
SQ	SEQUENCE 809 AA; 91022 MW; 9EA5287B268910E3 CRC64;

Query Match            4.2%; Score 178; DB 16; Length 809;  
 Best Local Similarity   23.2%; Pred. No. 2e+05;  
 Matches         72; Conservative      66; Mismatches   125; Indels   48; Gaps   14;

Qy	163	VNKVKKAKRIQE-WVATVSPAMIRITGVW----	LLKLNSFFWNIOIHKGQLEMYKAA	217
Dd	476	VTEIADYKPFRLDTHLIAPARYLVSGFGAGLPKI-----	CGCVKEVKAL	524
Qy	218	TETNLPLELPHVRHSHIDYLLLTIFILCHNIKAP-VIASGNNLNIPISFTLIHLGGFFI	276	
Dd	525	A-SRYDLYPTTHRSHLDSEVAEGIKWLGLVPVRYAADKKVWATPGLASVLKSLGAYMV	583	
Qy	277	RRRLEDTPGRKOVLYRALHLGHIVELLROOQFLIEFLGTRSRSGKTSCARAGLLSVVV	336	
Dd	584	DRKN-----RNILYLECLTQYSTWMLEAGIPTLVYPEGTSRSRTGGILPIKTGILTSTV	637	
Qy	337	DTLSTNVIPDLIIIPGISVDRIITECHNGEOLGPKPNESLSVARGVTRMLRKNYGCV	396	
Dd	638	EAYK-HTPGEVIVPVILSYENVEP-----DDEFCGDKKS-----	GFKDFFYKRKE-V	684
Qy	397	RVPDAQPFSKLKEYLESOSQKPVSAL-LSLEQA-----	LLPAILPSRPSDAADEGRDTSI	449
Dd	685	YMDLCPEIPVSRYYTHEED--PTSGIGFEITCGWKYRILPNQLVARM--	IIVSGGEVNT	740
Qy	450	NEGRNATDESIL	460	
Dd	741	NELRNLIKETL	751	

Search completed: January 10, 2004, 20:20:50  
Job time : 75 secs

**THIS PAGE BLANK (USPTO)**

Kapeller-Libermann R. Logan TJ:

XX WPI; 2002-280934/32.  
 DR N-PSDB; ABK11094.  
 XX New acyltransferase-1, specifically glycerol-3-phosphate  
 PT acyltransferase, nucleic acids and proteins, useful for treating or  
 PT diagnosing e.g. metabolic disorders, weight disorders and appetite  
 PT regulation disorders  
 XX Claim 13; Fig 1; 122pp; English.  
 XX The invention relates to acyltransferase (ACTR-1) nucleic acids and  
 CC proteins. These are useful as modulating agents in regulating a variety  
 CC of cellular processes (e.g. cellular metabolism, energy homeostasis,  
 CC and/or cellular proliferation, growth, differentiation and/or migration).  
 CC The nucleic acids, proteins, protein homologues, fragments, antibodies,  
 CC peptides, peptidomimetics and small molecules can be used in screening  
 CC assays, predictive medicine (e.g. diagnostic or prognostic assays,  
 CC monitoring clinical trials, and pharmacogenetics), and in methods of  
 CC treatment (e.g. therapeutic or prophylactic). The ACTR-1 polypeptide or  
 CC its fragments are useful as reagents or targets in assays for the  
 CC treatment and/or diagnosis of ACTR-1 mediated or related disorders, which  
 CC include metabolic disorders such as disorders of energy homeostasis e.g.  
 CC diabetes, impaired glucose tolerance, insulin resistance, hyperglycaemia,  
 CC hypercholesterolaemia, hyperlipoproteinaemia, hypertriglyceridaemia,  
 CC and/or hyperlipidaemia; diabetic complications including atherosclerosis,  
 CC stroke, retinopathy, nephropathy, and peripheral neuropathy; weight  
 CC disorders and appetite regulation disorders, e.g. obesity, cachexia,  
 CC anorexia and bulimia. Proteins may also be used to screen for naturally  
 CC occurring ACTR-1 substrates, to screen for drugs or compounds which  
 CC modulate ACTR-1 expression and to treat disorders characterised by  
 CC insufficient or excessive production of ACTR-1 protein. The nucleic acids  
 CC may be used as hybridisation probes or primers, to express ACTR-1  
 CC protein, to detect ACTR-1 mRNA or a genetic alteration in an ACTR-1 gene,  
 CC to modulate ACTR-1 activity, in chromosome mapping, tissue typing and  
 CC forensic biology, and as surrogate markers (e.g. markers for precursors  
 CC of disease states, for predisposition of disease states, of drug  
 CC activity or of the pharmacogenomic profile of a subject). Modulators of  
 CC ACTR-1 protein activity or nucleic acid expression may be used to treat  
 CC a subject having a cardiovascular disorder or a triglyceride metabolism  
 CC disorder characterised by aberrant ACTR-1 protein activity or nucleic  
 CC acid expression. Anti-ACTR-1 antibodies can be used to detect and  
 CC isolate ACTR-1 proteins, regulate the bioavailability of ACTR-1  
 CC proteins, and modulate ACTR-1 activity. The present sequence  
 CC represents the amino acid sequence of human ACTR-1.  
 XX  
 SQ Sequence 828 AA;

Query Match 100.0%; Score 4280; DB 23; Length 828;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSALTLGTIDVSYLPHSSEYSGVRCKTSEEWGCGPPTVFRSATLKWESLSRKR 60  
 DB 1 MDSALTLGTIDVSYLPHSSEYSGVRCKTSEEWGCGPPTVFRSATLKWESLSRKR 60

QY 61 PFVGRCCYCTPQSDWKFFNPSPSLGLRNVIIYNETHRHGLARRLSYVLFQERDV 120  
 DB 61 PFVGRCCYCTPQSDWKFFNPSPSLGLRNVIIYNETHRHGLARRLSYVLFQERDV 120

QY 121 HKGMFATNTVENVLSRVOEAEVAALNPPGSAQQQSKAVNKKVKKAKRILQEWAT 180  
 DB 121 HKGMFATNTVENVLSRVOEAEVAALNPPGSAQQQSKAVNKKVKKAKRILQEWAT 180

QY 181 VSPAMIRLTGWLKLLFNSFFNIQIHKGQLEVMKAAATETNPLLPVHRSHIDYLLT 240  
 DB 181 VSPAMIRLTGWLKLLFNSFFNIQIHKGQLEVMKAAATETNPLLPVHRSHIDYLLT 240

QY 241 FILFCHNIKAPYTAGNNINIPFSTLIHKLGGFFRRRLDETPDGKQVLYRALLHGI 300  
 DB 241 FILFCHNIKAPYTAGNNINIPFSTLIHKLGGFFRRRLDETPDGKQVLYRALLHGI 300

QY 301 VELLRQQQFLIEFLEGRSRSGKTSARAGLLSVVVDLTSTNVIPDILIPVGISYDRII 360

DB 301 VELLRQQQFLIEFLEGRSRSGKTSARAGLLSVVVDLTSTNVIPDILIPVGISYDRII 360  
 QY 361 EGHYNGEQLGKPKNESLWSVARGVIRMLRKNYGCVRVDFQAPFSIKVLESOSQKPVSA 420  
 DB 361 EGHYNGEQLGKPKNESLWSVARGVIRMLRKNYGCVRVDFQAPFSIKVLESOSQKPVSA 420  
 QY 421 LLSLQALLPALPSRPDAADDEGRDTSINESRNATDESRRRLTANLAELHILFTASKSC 480  
 DB 421 LLSLQALLPALPSRPDAADDEGRDTSINESRNATDESRRRLTANLAELHILFTASKSC 480  
 QY 481 AIMSTHIVACLLYRHRQIGDILSTLVEDFFVMKEVLARDPDLGFGSGNSEDVVMHAIQLL 540  
 DB 481 AIMSTHIVACLLYRHRQIGDILSTLVEDFFVMKEVLARDPDLGFGSGNSEDVVMHAIQLL 540  
 QY 541 GNCVITHTSRNDEFFITPSTTPVPSFELNFSYNGVLHVFMELIACSLYAVLNKRGIG 600  
 DB 541 GNCVITHTSRNDEFFITPSTTPVPSFELNFSYNGVLHVFMELIACSLYAVLNKRGIG 600  
 QY 601 GPTSTPPNLIISOQLVRKAASLCYLLSNEGTTSLPCQTFYQVCHETVGVKFIQVGLTVAE 660  
 DB 601 GPTSTPPNLIISOQLVRKAASLCYLLSNEGTTSLPCQTFYQVCHETVGVKFIQVGLTVAE 660  
 QY 661 HDDQEDISPLAEQWQDKKLPEPLSWRSDEDEDESDFGEEQDCYLKVSKSKEHQOFTIF 720  
 DB 661 HDDQEDISPLAEQWQDKKLPEPLSWRSDEDEDESDFGEEQDCYLKVSKSKEHQOFTIF 720  
 QY 721 LQRLGLPALLEAYSSAAIFVHNFSGVPPEPEYLQKLHKYLIITERNVAVYASATYCLVK 780  
 DB 721 LQRLGLPALLEAYSSAAIFVHNFSGVPPEPEYLQKLHKYLIITERNVAVYASATYCLVK 780  
 QY 781 NAVKMFKDIDGVFKETKQKRVSVLELSSSTFLPCNQKQLLEYILSPFWL 828  
 DB 781 NAVKMFKDIDGVFKETKQKRVSVLELSSSTFLPCNQKQLLEYILSPFWL 828

## RESULT 2

ABG66665  
 ID ABG66665 standard; Protein; 828 AA.  
 XX  
 AC ABG66665;  
 XX  
 DT 29-AUG-2002 (first entry)  
 XX  
 DE Human glycerol-3-phosphate acyltransferase hGPAT.  
 XX  
 KW Fatty acid regulated gene; polyunsaturated fatty acid disorder;  
 KW PUFA disorder; eczema; cardiovascular disorder; hypertriglyceridaemia;  
 KW dyslipidaemia; atherosclerosis; coronary artery disease;  
 KW cerebrovascular disease; peripheral vascular disease; inflammation;  
 KW sinusitis; asthma; pancreatitis; osteoarthritis; rheumatoid arthritis;  
 KW acne; body weight disorder; obesity; cachexia; anorexia;  
 KW psychiatric disorder; cancer; cystic fibrosis; pre-menstrual syndrome;  
 KW diabetes; diabetic complication; genetic polymorphism.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200240666-A2.  
 XX  
 PD 23-MAY-2002.  
 XX  
 PF 19-NOV-2001; 2001WO-CA01632.  
 XX  
 PR 17-NOV-2000; 2000US-248589P.  
 XX  
 PA (XENO-) XENON GENETICS INC.  
 XX  
 PI Winther MD, Goldberg YP, Knickle LC, Haardt M, Allen SJ, Ponton A;  
 PI De Antueno RJ, Jenkins DK, Nwaka SO;  
 XX WPI; 2002-508327/54.  
 DR N-PSDB; ABK94821.  
 XX

PT Novel isolated polypeptide segment encoded by fat regulated genes,  
PT useful for diagnosing the presence of or a predisposition for a  
PT disorder involving fatty acid regulated genes in a subject  
PS  
XX Claim 14; Fig 28; 225pp; English.  
CC The invention describes an isolated polypeptide segment (I) whose genes  
CC are fat regulated. (I) or the polynucleotide encoding it (II) are useful  
CC for diagnosing the presence of or a predisposition for a disorder  
CC involving fatty acid regulated genes in a subject. A composition  
CC containing (I) or (II) is useful for treating a disorder involving fatty  
CC acid regulated genes, where the disorder is selected from a  
CC polyunsaturated fatty acid (PUFA) disorder, eczema, cardiovascular  
CC disorders (such as hypertriglyceridaemia, dyslipidaemia, atherosclerosis,  
CC coronary artery disease, cerebrovascular disease or peripheral vascular  
CC disease), inflammation (such as sinusitis, asthma, pancreatitis,  
CC osteoarthritis, rheumatoid arthritis or acne), body weight disorders  
CC (such as obesity, cachexia or anorexia), psychiatric disorders, cancer,  
CC cystic fibrosis, pre-menstrual syndrome, diabetes, and diabetic  
CC complications. (I) or (II) is useful as research agent and materials for  
CC discovery of treatments and diagnostics for a disease, particularly human  
CC disease. (II) is useful for constructing nucleotide probes and primers,  
CC for detecting genetic polymorphism, for detecting changes in the level of  
CC expression of (II), and as a diagnostic tool. This is the amino acid  
CC sequence of a protein regulated by fatty acids.

SQ Sequence 828 AA;

Query Match 100.0%; Score 4279; DB 23; Length 828;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 827; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDESALTGTIDVSLPHSSYSVGRCKHTSEEWGECGFRPTFRSATLKWKESLMGRKR 60  
DB 1 MDESALTGTIDVSLPHSSYSVGRCKHTSEEWGECGFRPTFRSATLKWKESLMGRKR 60  
QY 61 PFVGRCCYCTPQSDWKDFNPSISGLRNVIYNETHRHGWLARLSVYLFQERDV 120  
DB 61 PFVGRCCYCTPQSDWKDFNPSISGLRNVIYNETHRHGWLARLSVYLFQERDV 120  
QY 121 HKGFATNTENVLNSSRVQSAIEVAELNPDGSAQQSKAVNKVKKAKRIILOEMVAT 180  
DB 121 HKGFATNTENVLNSSRVQSAIEVAELNPDGSAQQSKAVNKVKKAKRIILOEMVAT 180  
QY 181 VSPAMIRLTGWLLKLFNSFFWNIQHKGLEWKAATETNLPILFLPVHRSHIDYLLLT 240  
DB 181 VSPAMIRLTGWLLKLFNSFFWNIQHKGLEWKAATETNLPILFLPVHRSHIDYLLLT 240  
QY 241 FILFCHNIKAPYIASGNNLNIPFSTLIHKLGGFFIRRRDETDPGRKDVLYRALLHGH 300  
DB 241 FILFCHNIKAPYIASGNNLNIPFSTLIHKLGGFFIRRRDETDPGRKDVLYRALLHGH 300  
QY 301 VELLRQQQFLFIFLEGTSRSGKTSARAGLLSVVVDLTSTNVPDILIIIPVGSYDRII 360  
DB 301 VELLRQQQFLFIFLEGTSRSGKTSARAGLLSVVVDLTSTNVPDILIIIPVGSYDRII 360  
QY 361 EGHYNGEQLKPKKNESLWSVARGVIRMLRNKNGCVRDVDFAPPSLKEYLESQKPVSA 420  
DB 361 EGHYNGEQLKPKKNESLWSVARGVIRMLRNKNGCVRDVDFAPPSLKEYLESQKPVSA 420  
QY 421 LLSLEQALLPALPRSPDAADGDRDTSINERNATDESRLRLIANLAHILFTASKSC 480  
DB 421 LLSLEQALLPALPRSPDAADGDRDTSINERNATDESRLRLIANLAHILFTASKSC 480  
QY 481 AIMSTHIVACLLLYHRRQIDLTSTVEDFFVWKEEVLARDDFLGSNGSDVVMHAIQLL 540  
DB 481 AIMSTHIVACLLLYHRRQIDLTSTVEDFFVWKEEVLARDDFLGSNGSDVVMHAIQLL 540  
QY 541 GNCVTITHTSRNDEFFITPSTVPSVFNLYNGVLHVFIMEATIASCLYAVLNKRLG 600  
DB 541 GNCVTITHTSRNDEFFITPSTVPSVFNLYNGVLHVFIMEATIASCLYAVLNKRLG 600  
QY 601 GPTSTPPNLIISOQLVRKAASLCYLLSNEGTISLPCQTFYQVCHETVKGFIQYGLITVAE 660

DB 601 GPTSTPPNLIISOQLVRKAASLCYLLSNEGTISLPCQTFYQVCHETVKGFIQYGLITVAE 660  
QY 661 HDQEDISPSIAEQOWDKLPEPLSWRSDEDESDSDFGEQRDCYLKVSQSKHQQTTF 720  
DB 661 HDQEDISPSIAEQOWDKLPEPLSWRSDEDESDSDFGEQRDCYLKVSQSKHQQTTF 720  
QY 721 LQRLGLLEAYSAALFVHNFSGVPPEPYLQKLHKYLITRTRNVAVVAESATYCLVK 780  
DB 721 LQRLGLLEAYSAALFVHNFSGVPPEPYLQKLHKYLITRTRNVAVVAESATYCLVK 780  
QY 781 NAVKMFIDIGVFKETKOKRVSVLELSSTFLPCQNRQKLELYLSFVVL 828  
DB 781 NAVKMFIDIGVFKETKOKRVSVLELSSTFLPCQNRQKLELYLSFVVL 828  
RESULT 3  
AAE22144  
ID AAE22144 standard; Protein; 828 AA.  
XX  
AC AAE22144;  
XX  
DT 25-JUL-2002 (first entry)  
XX  
DB Human TRNFR-6 protein.  
XX  
KW Human; transferase; developmental disorder; cell proliferative disorder;  
KW TRNFR-6; neurological disorder; autoimmune disorder; parasitic infection;  
KW inflammatory disorder; endocrine; antiparasitic; immunosuppressive;  
KW cytostatic; neurological.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 172..197  
FT Domain /note= "Transmembrane domain"  
FT Domain 215..412  
FT Domain /note= "Acetyltransferase domain"  
XX  
FN WO200226950-A2.  
XX  
PD 04-APR-2002.  
XX  
PF 28-SEP-2001; 2001WO-US30424.  
XX  
PR 29-SEP-2000; 2000US-236523P.  
PR 06-OCT-2000; 2000US-238481P.  
PR 27-OCT-2000; 2000US-244025P.  
PR 03-NOV-2000; 2000US-246001P.  
PR 09-NOV-2000; 2000US-247931P.  
PR 16-NOV-2000; 2000US-249639P.  
PR 21-NOV-2000; 2000US-252819P.  
XX  
FA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Lal PG, Tang YT, Yue H, Burford N, Gandhi AR, Warren BA, Yao MG;  
PI Tribouley CM, Baughn MR, Lee EA, Hafalia AJA, Lu Y, Griffin JA;  
PI Sanjanwala MS, Ding L,  
XX  
DR WPI; 2002-362492/39.  
XX  
PS N-PSDB; AAD35221.  
PT Novel human transferase polypeptides and polynucleotides, useful in  
treating e.g., cell proliferative and autoimmune disorders -  
XX  
Claim 61; Page 126-128; 168pp; English.  
XX  
CC The present invention relates to novel human transferases (TRNFR) and  
CC polynucleotides encoding such proteins. The TRNFR proteins are useful  
CC for treating disorders associated with a decreased expression of  
CC functional TRNFR, e.g., cell proliferative, developmental, neurological,  
CC autoimmune/inflammatory disorders and parasitic infections. Antagonists  
CC of TRNFR proteins are useful in treating disorders associated with



CC increased activity of TRNFR. The present sequence is human TRNFR-6  
XX protein.  
SQ Sequence 828 AA;  
Query Match 100.0%; Score 4279; DB 23; Length 828;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 827; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDESALTGLTIDVSYLPHSSEYSGVCKHTSEWGECEGFPRTVFRSATLWKESLSMRKR 60  
DB 1 MDESALTGLTIDVSYLPHSSEYSGVCKHTSEWGECEGFPRTVFRSATLWKESLSMRKR 60  
QY 61 PFVGRCCYCTPQSDWDFNPSISGLRNVIYINETHRHGWLARLSYVLFIQERDV 120  
DB 61 PFVGRCCYCTPQSDWDFNPSISGLRNVIYINETHRHGWLARLSYVLFIQERDV 120  
QY 121 HKGMFATNTENVLNSSRVQEAIAEVAALNPDGSAQQQSKAVNKKAKRILOEMVAT 180  
DB 121 HKGMFATNTENVLNSSRVQEAIAEVAALNPDGSAQQQSKAVNKKAKRILOEMVAT 180  
QY 181 VSPAMIRLTGWLLKLFNSFFWNIQHKQLEWVKAATETNLPFLPVHRSHIDYLLLT 240  
DB 181 VSPAMIRLTGWLLKLFNSFFWNIQHKQLEWVKAATETNLPFLPVHRSHIDYLLLT 240  
QY 241 FILFCHNIKAPYIASGNNLNIPFSTLIHKLGFFIRRLDETDPGKQVLYRALLHGH 300  
DB 241 FILFCHNIKAPYIASGNNLNIPFSTLIHKLGFFIRRLDETDPGKQVLYRALLHGH 300  
QY 301 VELLROOQFLEIFLEGTRSRSGKTSARAGLLSVVDTLSTNVIPIIDILIPVGISYDR 360  
DB 301 VELLROOQFLEIFLEGTRSRSGKTSARAGLLSVVDTLSTNVIPIIDILIPVGISYDR 360  
QY 361 EGHNGEOLGPKKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLESQKPVSA 420  
DB 361 EGHNGEOLGPKKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLESQKPVSA 420  
QY 421 LLSLEQALLPAILPRSPDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
DB 421 LLSLEQALLPAILPRSPDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
QY 481 AIMSTHIVACLLYRHRQGDISTLVEDFFVMKEEVLARDFDLFGSGNSEDDVVMHAIQL 540  
DB 481 AIMSTHIVACLLYRHRQGDISTLVEDFFVMKEEVLARDFDLFGSGNSEDDVVMHAIQL 540  
QY 541 GNCVTIHTSRNDEFFITPTSTVSPFELNPNYNGVLHVFMIAIACSLYAVLNKRG 600  
DB 541 GNCVTIHTSRNDEFFITPTSTVSPFELNPNYNGVLHVFMIAIACSLYAVLNKRG 600  
QY 601 GPTSTPNLISQBLVRKAASLCYLLSNEGTSILPCOTFYQVCHETVKGFIQYGLTVAE 660  
DB 601 GPTSTPNLISQBLVRKAASLCYLLSNEGTSILPCOTFYQVCHETVKGFIQYGLTVAE 660  
QY 661 HDQEDISPSLAQWQDKLPELSWRSDEEDSDSFGGEQRDCYLKVSQKEHQFIF 720  
DB 661 HDQEDISPSLAQWQDKLPELSWRSDEEDSDSFGGEQRDCYLKVSQKEHQFIF 720  
QY 721 LQRLGLPLLEYSAAI FVNFSGPVEPEYLOKLHKYLI TRTRNVNVAESATYCLVK 780  
DB 721 LQRLGLPLLEYSAAI FVNFSGPVEPEYLOKLHKYLI TRTRNVNVAESATYCLVK 780  
QY 781 NAVMFKDIGVFKETKOKRVSVLELSTFLPCQNRKLLLEYILSFVVL 828  
DB 781 NAVMFKDIGVFKETKOKRVSVLELSTFLPCQNRKLLLEYILSFVVL 828

RESULT 4  
AA027073  
ID AA027073 standard; Protein; 828 AA.  
XX  
AC AA027073;  
AC  
DT 22-MAY-2003 (first entry)  
XX  
DT

GPAM related protein sequence, SEQ ID No 2.  
XX  
DE Antidiabetic; nephrotropic; neuroprotective; ophthalmological; human;  
XX mitochondrial sn-glycerol-3-phosphate acyltransferase; GPAM;  
KW diabetic complication; retinopathy; neuropathy; enzyme.  
KW  
OS Homo sapiens.  
XX  
PN W02003008590-A1.  
XX  
PD 30-JAN-2003.  
XX  
PF 16-JUL-2002; 2002WO-JP07189.  
XX  
PR 16-JUL-2001; 2001JP-0215337.  
XX  
PA (KISP) KISSEI PHARM CO LTD.  
XX  
PI Sakamoto S, Onota H, Sugano S, Nakamura Y;  
XX  
XX WPI; 2003-229583/22.  
DR N-PSDB; AAL55475.  
XX  
PT Human mitochondrial sn-glycerol-3-phosphate acyltransferase and  
XX antagonists for treatment and prevention of diabetic complications -  
PS Disclosure; Page 40-46; 56pp; Japanese.  
XX  
CC The invention relates to a novel protein having human mitochondrial sn-  
CC glycerol-3-phosphate acyltransferase (GPAM) activity. The novel protein  
CC with GPAM activity can be used in the prevention and treatment of  
CC diabetic complications, including retinopathy and neuropathy, by  
CC administration of antagonists to human GPAM. This sequence represents a  
CC human protein relating to the GPAM activity protein of the invention.  
SQ Sequence 828 AA;  
Query Match 99.9%; Score 4275; DB 24; Length 828;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MDESALTGLTIDVSYLPHSSEYSGVCKHTSEWGECEGFPRTVFRSATLWKESLSMRKR 60  
DB 1 MDESALTGLTIDVSYLPHSSEYSGVCKHTSEWGECEGFPRTVFRSATLWKESLSMRKR 60  
QY 61 PFVGRCCYCTPQSDWDFNPSISGLRNVIYINETHRHGWLARLSYVLFIQERDV 120  
DB 61 PFVGRCCYCTPQSDWDFNPSISGLRNVIYINETHRHGWLARLSYVLFIQERDV 120  
QY 121 HKGMFATNTENVLNSSRVQEAIAEVAALNPDGSAQQQSKAVNKKAKRILOEMVAT 180  
DB 121 HKGMFATNTENVLNSSRVQEAIAEVAALNPDGSAQQQSKAVNKKAKRILOEMVAT 180  
QY 181 VSPAMIRLTGWLLKLFNSFFWNIQHKQLEWVKAATETNLPFLPVHRSHIDYLLLT 240  
DB 181 VSPAMIRLTGWLLKLFNSFFWNIQHKQLEWVKAATETNLPFLPVHRSHIDYLLLT 240  
QY 241 FILFCHNIKAPYIASGNNLNIPFSTLIHKLGFFIRRLDETDPGKQVLYRALLHGH 300  
DB 241 FILFCHNIKAPYIASGNNLNIPFSTLIHKLGFFIRRLDETDPGKQVLYRALLHGH 300  
QY 301 VELLROOQFLEIFLEGTRSRSGKTSARAGLLSVVDTLSTNVIPIIDILIPVGISYDR 360  
DB 301 VELLROOQFLEIFLEGTRSRSGKTSARAGLLSVVDTLSTNVIPIIDILIPVGISYDR 360  
QY 361 EGHNGEOLGPKKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLESQKPVSA 420  
DB 361 EGHNGEOLGPKKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLESQKPVSA 420  
QY 421 LLSLEQALLPAILPRSPDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
DB 421 LLSLEQALLPAILPRSPDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480

QY 481 AIMSTHIVACLLLYRHROGIDILSTLVEDFFVMKEEVLARDFDLGFSGNSEDDVWHAIQLL 540  
DB 481 AIMSTHIVACLLLYRHROGIDILSTLVEDFFVMKEEVLARDFDLGFSGNSEDDVWHAIQLL 540  
QY 541 GNCVTITHTSRNDEFFITPSTTPVPSVFEINFNYSNGVLHVFMEALIIACSLYAVLNKRGGLG 600  
DB 541 GNCVTITHTSRNDEFFITPSTTPVPSVFEINFNYSNGVLHVFMEALIIACSLYAVLNKRGGLG 600  
QY 601 GPTSTPPNLIISOQLVRKAASICYLLSNEGTSILPCQTFYQVCHETVGVKFIQYGLITVAE 660  
DB 601 GPTSTPPNLIISOQLVRKAASICYLLSNEGTSILPCQTFYQVCHETVGVKFIQYGLITVAE 660  
QY 661 HDDQEDISPSLAEOQWKKLPEPLSWRDEDEDDSDGFEORDCYLVKVSQKHEHQOFTTF 720  
DB 661 HDDQEDISPSLAEOQWKKLPEPLSWRDEDEDDSDGFEORDCYLVKVSQKHEHQOFTTF 720  
QY 721 LQRLIGPLLEAYSSAAIIVHNFSGVPPEYVLOKLHKYLIIRTRNNVAVYAESATYCLVVK 780  
DB 721 LQRLIGPLLEAYSSAAIIVHNFSGVPPEYVLOKLHKYLIIRTRNNVAVYAESATYCLVVK 780  
QY 781 NAVRMFKDIGVFKETKQKRVSVLELSSTFLPQCNRQKLLLEYILSPVVL 828  
DB 781 NAVRMFKDIGVFKETKQKRVSVLELSSTFLPQCNRQKLLLEYILSPVVL 828

RESULT 5

ID ABUS4605  
XX ABUS4605 standard; Protein; 828 AA.

AC ABUS4605;

DT 03-JUN-2003 (first entry)

DE Human NOVX polypeptide #64.

XX Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;  
KW hypertension; congenital heart defect; aortic stenosis; valve disease;  
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;  
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;  
KW tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;  
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;  
KW Parkinson's disease; immune disorder; haematopoietic disorder;  
KW haemophilia; hypercoagulation; Crohn's disease; cancer.

OS Homo sapiens.

XX WO200281498-A2.

XX 17-OCT-2002

XX 03-APR-2002; 2002WO-US10780.

XX 03-APR-2001; 2001US-281086P.  
PR 03-APR-2001; 2001US-281136P.  
PR 05-APR-2001; 2001US-281863P.  
PR 06-APR-2001; 2001US-281906P.  
PR 06-APR-2001; 2001US-282020P.  
PR 10-APR-2001; 2001US-282930P.  
PR 10-APR-2001; 2001US-282934P.  
PR 12-APR-2001; 2001US-283512P.  
PR 13-APR-2001; 2001US-283710P.  
PR 17-APR-2001; 2001US-284234P.  
PR 19-APR-2001; 2001US-285325P.  
PR 20-APR-2001; 2001US-285381P.  
PR 20-APR-2001; 2001US-285609P.  
PR 23-APR-2001; 2001US-285748P.  
PR 23-APR-2001; 2001US-285890P.  
PR 24-APR-2001; 2001US-286068P.  
PR 25-APR-2001; 2001US-286292P.  
PR 27-APR-2001; 2001US-287213P.  
PR 02-MAY-2001; 2001US-288257P.  
PR 29-MAY-2001; 2001US-294164P.

PR 30-MAY-2001; 2001US-294484P.  
PR 18-JUN-2001; 2001US-298952P.  
PR 19-JUN-2001; 2001US-299237P.  
PR 19-JUN-2001; 2001US-299276P.  
PR 12-SEP-2001; 2001US-318750P.  
PR 25-SEP-2001; 2001US-324800P.  
PR 25-SEP-2001; 2001US-324802P.  
PR 27-SEP-2001; 2001US-325684P.  
PR 17-OCT-2001; 2001US-330143P.  
PR 14-NOV-2001; 2001US-332131P.  
PR 14-NOV-2001; 2001US-332240P.  
PR 14-NOV-2001; 2001US-332779P.  
PR 21-NOV-2001; 2001US-332115P.  
PR 04-DEC-2001; 2001US-337621P.  
PR 03-JAN-2002; 2002US-345783P.  
PR 16-JAN-2002; 2002US-350251P.  
PR 02-APR-2002; 2002US-0114270.

(CURA-) CURAGEN CORP.

Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA, Patturajan M;  
Liu X, Gusev VY, Li L, Vernet CAM, Zerhusen BD, Gorman L;  
Shenoy SG, Pena CEA, Smithson G, Burgess CE, Gerlach V;  
Padigaru M, Shinkets RA, Gangolli EA, Raupier RJ, Casman SJ, Ji W;  
Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;  
MacDougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;  
Ellerman K;

WPI; 2003-046858/04.  
N-PSDB; ABX72233.

New isolated NOVX polypeptide useful for treating atherosclerosis,  
metabolic disorders, diabetes, obesity, infectious disease, anorexia,  
neurodegenerative disorders, Alzheimer's disease and cancer

Claim 1; Page 232; 666pp; English.

The invention relates to human polypeptides, termed NOVX, and the  
polynucleotides encoding them. The polypeptides and polynucleotides are  
useful for diagnosing disease, and screening for potential therapeutic  
agents. The sequences are useful for treating metabolic disorders,  
cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic  
stenosis, atrial septal defect (ASD), atrioventricular canal defect,  
ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular  
septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,  
atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative  
disorders, Alzheimer's disease, Parkinson's disease, immune disorders,  
haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease  
and cancer. Sequences ABUS4542-ABUS4647 represent human NOVX polypeptides  
of the invention.

Sequence 828 AA;

Query Match 99.8%; Score 4270; DB 24; Length 828;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 826; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDESALTGLTIDVSYLPHSSEYSVGRCKHTSEWEGCGFRPTVPSATLKWESLMSRKR 60  
DB 1 MDESALTGLTIDVSYLPHSSEYSVGRCKHTSEWEGCGFRPTVPSATLKWESLMSRKR 60  
Qy 61 PFVGRCCYCTPQSWDKFFNPSIPSLGRNVIYINETHTRHGWLARLSVYLFQIQRDV 120  
DB 61 PFVGRCCYCTPQSWDKFFNPSIPSLGRNVIYINETHTRHGWLARLSVYLFQIQRDV 120  
Qy 121 HKGMFATNVTENVLNSRVQEAIAEVAALNPDGSAQQSKAVNKKVKKAKRIQLQEMVAT 180  
DB 121 HKGMFATNVTENVLNSRVQEAIAEVAALNPDGSAQQSKAVNKKVKKAKRIQLQEMVAT 180  
Qy 181 VSPAMIRLTGWLKLFNSFFWNIQIHKGQLEVMVKAATETNLLPLFVHRSHDYLILLT 240  
DB 181 VSPAMIRLTGWLKLFNSFFWNIQIHKGQLEVMVKAATETNLLPLFVHRSHDYLILLT 240

Qy 241 FILFCHNIKAPIYASGNLNPIFSTLIHKLGGFFIRRLDETDPGRKDVLYRALLHGI 300  
Db 241 FILFCHNIKAPIYASGNLNPIFSTLIHKLGGFFIRRLDETDPGRKDVLYRALLHGI 300  
Qy 301 VELLROQOFLEIFLEGTRSGKTSARAGLLSVVVDLTSTNVPDILLIPVGISYDRII 360  
Db 301 VELLROQOFLEIFLEGTRSGKTSARAGLLSVVVDLTSTNVPDILLIPVGISYDRII 360  
Qy 361 EGHVNGQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQKPVSA 420  
Db 361 EGHVNGQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQKPVSA 420  
Qy 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSC 480  
Db 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSC 480  
Qy 481 AIMSTHIVACLLLYRHRQIDLSLTVDFVFMKEVLARDFDLGFGSNGSDVVMHAIQLL 540  
Db 481 AIMSTHIVACLLLYRHRQIDLSLTVDFVFMKEVLARDFDLGFGSNGSDVVMHAIQLL 540  
Qy 541 GNCVTIHTSRNDEFFITPSTVPSVFEINFYNSGVLMVIMEAIACSLYAVLNKRGILG 600  
Db 541 GNCVTIHTSRNDEFFITPSTVPSVFEINFYNSGVLMVIMEAIACSLYAVLNKRGILG 600  
Qy 601 GPTSTPNLISQEQVRAKASLCVLLSNEGTSISLPCQTFVQVCHETVGKTIQYGLTVAE 660  
Db 601 GPTSTPNLISQEQVRAKASLCVLLSNEGTSISLPCQTFVQVCHETVGKTIQYGLTVAE 660  
Qy 661 HDQEDISPSLAQWKKLPEPLSWRSDEDESDSGEQRDCYLKVSQKSHQOPIIF 720  
Db 661 HDQEDISPSLAQWKKLPEPLSWRSDEDESDSGEQRDCYLKVSQKSHQOPIIF 720  
Qy 721 LQRLGPLELAYSAAIFVHNFSGPVPEPEYLOKLHKYLIITRERNVAVAESATYCLVK 780  
Db 721 LQRLGPLELAYSAAIFVHNFSGPVPEPEYLOKLHKYLIITRERNVAVAESATYCLVK 780  
Qy 781 NAVKMFKDIGVFKETKQKRVSVLELSSTFLPQCNQKLLLEYILSFVVL 828  
Db 781 NAVKMFKDIGVFKETKQKRVSVLELSSTFLPQCNQKLLLEYILSFVVL 828

RESULT 6  
ID ABP69808  
XX  
XX  
AC ABP69808;  
XX  
XX  
XX 20-JAN-2003 (first entry)  
XX Human polypeptide SEQ ID NO 1855.  
XX Human; genome mapping; gene therapy; food supplement; virus; fungus;  
KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
KW arthritis; cytostatic; immunomodulator; neotropic; neuroprotective;  
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
KW haemostatic; vulnary; fungicide; antibacterial; virucide; protozoacide;  
KW antiarthritic.  
XX  
XX Homo sapiens.  
XX  
XX WO200270539-A2.  
XX  
XX 12-SEP-2002.  
XX  
XX 05-MAR-2002; 2002WO-US05095.  
XX  
XX 05-MAR-2001; 2001US-0799451.  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;

PI 102 FILFCHNIKAPIYASGNLNPIFSTLIHKLGGFFIRRLDETDPGRKDVLYRALLHGI 300  
PI 102 FILFCHNIKAPIYASGNLNPIFSTLIHKLGGFFIRRLDETDPGRKDVLYRALLHGI 300  
XX  
DR WPI; 2002-759812/82.  
DR N-PSDB; ABZ12025.  
XX  
XX New polynucleotides comprising sequences assembled from expressed  
PT sequence tags (ESTs), useful for treating cell-proliferative,  
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or  
PT platelet or coagulation disorders -  
XX  
PS Claim 9; SEQ ID NO 1855; 1012pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated polynucleotide (I) comprising a  
CC nucleotide sequence selected from any of 948 sequences  
CC (ABZ1119-ABZ12066) or their mature protein coding portion, active domain  
CC coding protein or complementary sequences. The polynucleotides are useful  
CC for identifying expressed genes or for physical mapping of human genome.  
CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular  
CC weight markers, as a food supplement, for generating antibodies, in  
CC medical imaging, screening and diagnostic assays and for treating  
CC cell-proliferative disorders (cancer), neurodegenerative diseases  
CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple  
CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid  
CC disorders, platelet or coagulation disorders, wound, burns, incision,  
CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,  
CC parasitic), arthritis, etc.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 705 AA;

Query Match 84.2%; Score 3602; DB 23; Length 705;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 124 MFATNTVNTVNLSSRRVOEAEVAAELNPDGSAQQSKAVNKVKKAKRILQBMVATVSP 183  
Db 1 MFATNTVNTVNLSSRRVOEAEVAAELNPDGSAQQSKAVNKVKKAKRILQBMVATVSP 60  
Qy 184 AMIRLTGWVLLKLFNSFFWNQIHKQLEVMKAAETNLPFLPVRSHIDYLLTFFIL 243  
Db 61 AMIRLTGWVLLKLFNSFFWNQIHKQLEVMKAAETNLPFLPVRSHIDYLLTFFIL 120  
Qy 244 FCHNIKAPIYASGNLNPIFSTLIHKLGGFFIRRLDETDPGRKDVLYRALLHGI 303  
Db 121 FCHNIKAPIYASGNLNPIFSTLIHKLGGFFIRRLDETDPGRKDVLYRALLHGI 180  
Qy 304 LROQOQFLEIFLEGTRSGKTSARAGLLSVVVDLTSTNVPDILLIPVGISYDRII 363  
Db 181 LROQOQFLEIFLEGTRSGKTSARAGLLSVVVDLTSTNVPDILLIPVGISYDRII 240  
Qy 364 YNGEQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQKPVSA 423  
Db 241 YNGEQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQKPVSA 300  
Qy 424 LEOALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSC 483  
Db 301 LEOALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSC 360  
Qy 484 STHIVACLLLYRHRQIDLSLTVDFVFMKEVLARDFDLGFGSNGSDVVMHAIQLL 543  
Db 361 STHIVACLLLYRHRQIDLSLTVDFVFMKEVLARDFDLGFGSNGSDVVMHAIQLL 420  
Qy 544 VTIHTSRNDEFFITPSTVPSVFEINFYNSGVLMVIMEAIACSLYAVLNKRGILG 603  
Db 421 VTIHTSRNDEFFITPSTVPSVFEINFYNSGVLMVIMEAIACSLYAVLNKRGILG 480  
Qy 604 STPPNLISQEQVRAKASLCVLLSNEGTSISLPCQTFVQVCHETVGKTIQYGLTVA 663  
Db 481 STPPNLISQEQVRAKASLCVLLSNEGTSISLPCQTFVQVCHETVGKTIQYGLTVA 540

QY 664 QEDISPLAEQWKKLPEPLSWRSDEDEDSDFGEORDCYLKVSQSKHQQTITFLOR 723  
 DB 541 QEDISPLAEQWKKLPEPLSWRSDEDEDSDFGEORDCYLKVSQSKHQQTITFLOR 600  
 QY 724 LLGPLLEAYSSAAIFVHNFSGPVPEYLOKLHKYLIITRTNNAVVAESATYCLVKNAV 783  
 DB 601 LLGPLLEAYSSAAIFVHNFSGPVPEYLOKLHKYLIITRTNNAVVAESATYCLVKNAV 660  
 QY 784 KMFKDGVFKETQKRVSVLELSSTFLPQCNRQKLLEYILSFVVL 828  
 DB 661 KMFKDGVFKETQKRVSVLELSSTFLPQCNRQKLLEYILSFVVL 705

RESULT 7  
 AAE22145 ID AAE22145 standard; Protein; 801 AA.  
 XX AAE22145;  
 AC AAE22145;

DT 25-JUL-2002 (first entry)  
 DE Human TRNFR-7 protein.

KW Human; transferrase; developmental disorder; cell proliferative disorder;  
 KW TRNFR-7; neurological disorder; autoimmune disorder; parasitic infection;  
 KW inflammatory disorder; endocrine; antiparasitic; immunosuppressive;  
 KW cytostatic; neurological.

OS Homo sapiens.  
 XX WO200226950-A2.  
 PN 04-APR-2002.

PD 28-SEP-2001; 2001WO-US30424.

PF 29-SEP-2000; 2000US-236523P.  
 PR 06-OCT-2000; 2000US-238481P.  
 PR 27-OCT-2000; 2000US-244025P.  
 PR 03-NOV-2000; 2000US-246001P.  
 PR 09-NOV-2000; 2000US-247931P.  
 PR 16-NOV-2000; 2000US-249639P.  
 PR 21-NOV-2000; 2000US-252819P.

XX (INCY-) INCYTE GENOMICS INC.

XX Lal PG, Tang YT, Yue H, Burford N, Gandhi AR, Warren BA, Yao MG;  
 PI Tribouley CM, Baughn MR, Lee EA, Hafalia AJA, Lu Y, Griffin JA;  
 PI Sanjanwala MS, Ding L;

XX WPI; 2002-362492/39.  
 XX N-PSDB; AAD35222.

XX Novel human transferase polypeptides and polynucleotides, useful in  
 PT treating e.g., cell proliferative and autoimmune disorders -

XX Claim 62; Page 128-130; 168pp; English.

XX The present invention relates to novel human transferases (TRNFR) and  
 CC polynucleotides encoding such proteins. The TRNFR proteins are useful  
 CC for treating disorders associated with a decreased expression of  
 CC functional TRNFR, e.g., cell proliferative, developmental, neurological,  
 CC autoimmune/inflammatory disorders and parasitic infections. Antagonists  
 CC of TRNFR proteins are useful in treating disorders associated with  
 CC increased activity of TRNFR. The present sequence is human TRNFR-7  
 CC protein.

XX Sequence 801 AA;

XX Query Match 23.3%; Score 996.5; DB 23; Length 801;  
 XX Best Local Similarity 33.0%; Pred. No. 1.4e-92;  
 XX Matches 273; Conservative 146; Mismatches 348; Indels 61; Gaps 24;

QY 25 GRCKHTSEMGECGRFPTVFRSA--TLKWKESLMSRRKPPFVGRCCVCTPQSDWDFKFP 81  
 DB 7 GRCC-TOPRSPSGREASLWSSGFMKLEAVTTPGLKYPFVGRCCCTCTPKSWESLPHR 65  
 QY 82 SIPSLGNVTIYINETHTRHGMARLSVYLPFOERDVHKMFATNVTNVLSSRVOE 141  
 DB 66 SITDLGFCNVTLVKEENTFRGWLVRRLCYFLWLEQHIPP--CQDPQKIMESTGVON 122  
 QY 142 AIAEVAEALNPDGSAQQSKAVNKVKKAKRILQEMVATVSPAMIRLTGWLLKLFNSFF 201  
 DB 123 LLSGRV---PGTGEGQ--VPDLVKKVEQRILGHIOAPPFPFLVRLFSWALLPFLNCLF 176  
 QY 202 WNIQIHKGQEMVKAATETNLPPLFLPVHRSHIDYLLTLFLFCHNTKAPYIAGNNLNI 261  
 DB 177 LNVQLHKQMKVQKAAQAGLPLVLLSTHKLTLGILLPFMLLSQGLGVLRVWDSRACS 236  
 QY 262 PIFSTLIHKLGCFIRRLDETDPGRKDVLYRALLHGHIVELLRQQQFLEFLEGTSRS 321  
 DB 237 PALRALLRKLGLFLPPEASLSLDSSEGLLARAVVQAVIEQLLVSGQPLFLIFLEPPGAL 296  
 QY 322 G-KTSCARAGLLSVVVDTLSTNVIPIIIPVIGISYDRIIEGHYNGEQLGKPKNESLWS 380  
 DB 297 GPRLSALGOAVGVFVQVQVIGIVPDALLVPVAVTYDLVPDAPCDIDHASAPL---GLWT 353  
 QY 381 VARGVIRMLRNKYG-----CVRVDFAPFSLKYLESSQSKPVSALLSLEQALLPAILP- 434  
 DB 354 GALAVLASLWSRWCGRSHRCSRHLAQPFSLQEVIVS-ARSCWGRQTEQLLOQIVLGQ 412  
 QY 435 -SRPSDAADREGDTISINESRNATDESRLRLIANLAEHILFTASKSCAINSTHIVACLLL 493  
 DB 413 CTAVPDTKEQEWTPITGPLLALKEE-DQLLVRLSCHVLSASVGSASVMTAIMATLL 471  
 QY 494 YRHQIGIDLSTLVEDDFVMEKEVLARDFDLGCNSDVMVHAIQLLNCVITITHTSRND 553  
 DB 472 FKQKGVFLSOLLGEFSLWTEILLRGFDVGFSQLSLQHSLSLRAHVALLRI-RQG 530  
 QY 554 EFFITPTTVPVSEFELNFGVNLHVFIMEAIIACSLYAVLNKRLGLGQPTTTPN----- 608  
 DB 531 DLLVVPQPG-PGLTHLAQLSAELLPVFLSEAVGACAVRGLLAGR-----VPQGPWEL 582  
 QY 609 ----LISQEQVLKKAASLYLLSNEGHSISLPCQTFYQVCHETVGKFTQYGLITVAEHDDQ 664  
 DB 583 QGILLSSQNELYRQILLMHLPPQDLLLKPQSSYCYCQEVLDRLIQCGLL-VAEETPG 641  
 QY 665 EDISPLAEQWKKLPEPLSWR--SDEEDDS-DFGEORDCYLKVSQSKHQQTITFL 721  
 DB 642 SRPACDTGRQLSRK----LLWPSGDFTDSDDFGEAD-GRYFRLSQSHCHDFDLFL 696  
 QY 722 QRLGLPILLEAYSSAAIFVHNFSGPVPEPE--YLQKLHKYL-ITRTNNAVVAESATYCL 778  
 DB 697 CRLLSPLLKAFQAQAAFLR--QGQLPDTLGYTEQLFQFLQATAQESGI---FECAADPKL 751  
 QY 779 VKNAVVKFOIGVPKQKRVSVLELSSTFLPQCNRQKLLEYILSFV 826  
 DB 752 AISAVVTRDLGVLOQTTPSPAGPRLHLSPTFASLDNQEKLEQFIRQFI 799

RESULT 8  
 ABG32429 ID ABG32429 standard; Protein; 775 AA.  
 XX ABG32429;  
 AC ABG32429;  
 XX 29-NOV-2002 (first entry)  
 DT Human acyltransferase family member, 60491.  
 DE Human; acyltransferase; 60491; brain disorder; neuron; glia; astrocyte;  
 KW oligodendrocyte; cerebral oedema; hypoxia; ischaemia; hypotension;  
 KW hypoperfusion; intracranial haemorrhage; acute meningitis; tumour;  
 KW Alzheimer's disease; Pick's disease; tumour; acquired metabolic disease;  
 KW hypoglycaemia; hyperglycaemia; metabolism; leukodystrophy; pain;  
 KW Krabbe disease; inflammation; infection; ischaemia; arthritis; migraine;



RESULT 10  
ABG23092  
ID ABG23092 standard; Protein; 807 AA.  
XX AC ABG23092;  
XX AC  
XX DT 18-FEB-2002 (first entry)  
XX DT  
XX DE Novel human diagnostic protein #23083

Human: chromosome mapping: gene therapy: forensic:

XX food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX

PN  
WO200175067-A2.  
XX

PD 11-OCT-2001.  
XX

30-MAR-2001; 2001WO-US08631.  
PF  
XX

PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000. 2000US-0649167

XX  
PA (HVSE-) HVSECO INC

XX PT Dymanac PT Liu C Tang yf.

XX  
DP WPT: 2001-639362/73

DR N-PSDB; AAS87279.  
XX

PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics for forensic gene mapping identification of mutations

PT responsible for genetic disorders or other traits and to assess biodiversity -

PS Claim 20: SEQ ID NO 53451: 103nn. Encl 1 ab

CC The invention relates to isolated polynucleotide (T) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes, CC polymerase chain reaction (pcr) primers, oligomers and for chromosome

CC and gene mapping, and in recombinant production of (II). The CC polynucleotides are also used in diagnostics as expressed recombinant tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving

(II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue or molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II) (I) and (II) are useful for treating

disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity.

CC and to produce other types of data and products dependent on DNA and amino acid sequences. ABC00010-ABC20277 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC - specification, but was obtained in electronic format directly from WIPO  
CC at [ftn.wipo.int/pub/publ](http://ftn.wipo.int/pub/publ) not commences

XX  
SC  
Sequence 807 AA.

Query Match

Best Local Similarity	29.0%;	Pred: No. 1.2e-73;
Matches 233: Conservative	125:	Mismatches 289:
		Indels 145: Cons 20

56 MSPKPRFV/CBCCVUSCTBOSWPKRPNRST RST CI RMLTVINERTUTRNRCSWT L RBY CVJH ET 115

107 ICKVBBVCBCCTCTDVCBCEI EUPCETDI CEGNUTY IYPPENT

0x 116 0E9D74HKCMEATNPTEMT NSSBIVAEATAEVAAPENBDCEAOOCCKAIAINVTZVVKAVBTIO 17E

Db	151	-----	150
Qy	176	EMVATSPAMIRLTGTVLLKLFNSFFWNIOIHKGOLEMVKAAT-----ETN	221
Db	151	-----RUFSWALLRVLLKLLFUNVQLHKGOMVOKAAQQVQGVQTSUWLSPQEG	199
Qy	222	LPLFLFVHRSHIDYLLLTILFELCHNIKAPYIASGNNLNIPFSTLIHLKGGFFIRRLD	281
Db	200	LPLVLLSTHKTLDGILLPMLLSQGLGVLRVAVDSRACSPALRALLKGLFLPPEAS	259
Qy	282	ETPDGRKDVLYRALLHGHIVELLRQQQFLFIEGTRSRG-KTSCARAGLLSVVVDTLS	340
Db	260	LSLDSSEGLLARAVQAVIEQLLVSGQPLLIFLEPPGALGPLRSALGQAVGVFVQAVQ	319
Qy	341	TNVPIDILLIPVGISVDRIEYHNGEQLKPKKNESLMSVARGVIRMLRKNGV-----C	395
Db	320	VGIVPALLVPVAVTTDLVPDPCDDHASAPL---GLWTGALAVLRSLWSRWGCSHRIC	376
Qy	396	VRVDFAPQFSLKEYLESQSKPVSAALLSLQALLPAILP--SRPSDAADGRDTSINESR	453
Db	377	SRVHLAQPFSLQEVIVS-ARSCWGRQTLEQLLPQIVLGQCTAVPDTEKEQWTPITGPL	435
Qy	454	NATESLRRRLIANLAEHLIFTASKSCAINTHVACLLYHRHQGIDISTLVEDFFVMK	513
Db	436	LALKEE-DQLLVRLSCHVLSASVGSVAVMSTAIMATLLLFKHQKGVFLSQLLGEFSWLT	494
Qy	514	EEVLARDFDLFGSNGSEDEVVMAIQLLGNCTVITHTSRNDEFFITPTTVPSPFELNFS	573
Db	495	EEILVRGDFVFGSQLRSLLOHLSLRAHVALLRI-ROGDLVVPQPG-PGLTHLAQUS	552
Qy	574	NGVLHVFIMEATACSLIYAVLNKRGILGGPTSTPPN-----LISQEOLVKKAASLCY	624
Db	553	AELLVFLSAVACAVRGLLAGR-----VPPQGPWELQGLILLSQNELYQIILLMH	605
Qy	625	LLSNEGTSILPQCTFYOVCHETVKGFTIQYGLITVAEHDDQEDISPLAQDQWKKLPEPL	684
Db	606	LLPQDLLLLKPCQSSCYCQEVLDRLITQCLL-VABETPGSRPACDTGRLSRK----L	660
Qy	685	SWR--SDEEEDS-DEGEORDCYLVKSQSEHQQFIFLQRLGLLPLEAYSAAAFVHN	741
Db	661	LWPSGDFDSDSDDFEAD-GRYFRLSQSHGCFDFFLCRLSLPLLKAFQAAPFLR-718	
Qy	742	FSGVPPEPE--YLOKLHKYILITERNVAVYAESATVCLVKNVAKFKDGVKPKQKR	799
Db	719	-OQLPDTLGYTEQLFQFL-QATAEGEIFECAPKASPSAVWTFRLDGLVLQOTRSPA	776
Qy	800	VSVELSSTFLPQCNKQKLEYI	822
Db	777	GPRHLGLPLLLPALDNOEKLRTV	799
RESULT 11			
ID ABB11171			
AC	XX	ABB11171 standard; peptide; 156 AA.	
AC	XX	ABB11171;	
XX	XX		
DT	11-JAN-2002	(first entry)	
XX	XX	Human sn-glycerol-3-P acyltransferase homologue, SEQ ID NO:1541.	
DE	XX	Human; cytokine; cell proliferation; cell differentiation; growth factor;	
DE	XX	haematopoiesis regulation; tissue growth; immunomodulator; activin;	
KW	KW	inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;	
KW	KW	proliferation; metastasis; cancer; tumour; haematopoietic disorder;	
KW	KW	myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;	
KW	KW	chronic inflammatory condition; proliferative retinopathy;	
KW	KW	atherosclerosis; coronary heart disease; arterial ischaemia;	
KW	KW	bone disorder; osteoporosis; vascular growth disorder;	
KW	KW	tissue regeneration; wound healing; infection; immune disorder;	
KW	KW	cell culture; drug screening; gene therapy; antiinflammatory;	
KW	KW	antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;	
KW	KW	cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;	

KW	antifungal; vulnery; antiulcer.
XX	Homo sapiens.
XX	WO200157188-A2.
PN	09-AUG-2001.
PD	05-FEB-2001; 2001WO-US03800.
XX	03-FEB-2000; 2000US-0496914.
PR	27-APR-2000; 2000US-0560875.
XX	(HYSB-) HYSEQ INC.
PA	Tang YT, Liu C, Drmanac RT;
XX	WPI; 2001-457740/49.
XX	N-PSDB; ABA08415.
DR	Human proteins and DNA encoding sequences useful for preventing,
DR	treating or ameliorating a medical condition in a mammalian subject
PT	e.g. arthritis and cancer -
PT	Claim 20; Page 152; 1963pp; English.
XX	Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC	sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC	invention also relates to vectors and recombinant host cells comprising a
CC	nucleotide of the invention, methods of producing the novel polypeptides,
CC	antibodies against the polypeptides, methods of detecting the nucleotides
CC	or polypeptides in a sample, and methods of identifying compounds which
CC	bind to polypeptides of the invention. Although novel, many of the
CC	polypeptides of the invention have homology to known proteins, thereby
CC	giving an insight into their probable biological activities, and hence
CC	potential therapeutic applications. The polypeptides of the invention may
CC	have various activities, including cytokine, cell proliferation or cell
CC	differentiation activities; stem cell growth factor activity;
CC	haematopoiesis regulatory activity; tissue growth activity;
CC	immunomodulatory activity; activin- or inhibin-related activities;
CC	chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC	thrombolytic activities; receptor or ligand activities; or may be
CC	involved in oncogenesis, cancer cell proliferation or metastasis.
CC	Depending on their biological activities, polypeptides and nucleotides of
CC	the invention are useful for preventing, treating or ameliorating medical
CC	conditions, e.g., by protein or gene therapy. Such conditions include
CC	cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC	disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC	proliferative retinopathy, atherosclerosis, coronary heart disease,
CC	arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC	vascular growth. Polypeptides involved with tissue regeneration and
CC	repair (or nucleic acids encoding them) may be used to promote wound
CC	healing (e.g., of burns, incisions and ulcers), while those with
CC	immunomodulatory activities may be used in the treatment of viral,
CC	bacterial and fungal infections in addition to immune disorders.
CC	Polypeptides with growth factor activity may be used in cell cultures to
CC	promote cell growth. For example, such polypeptides may be used to
CC	manipulate stem cells in culture to give rise to neuroepithelial cells
CC	that can be used to augment or replace cells damaged by illness,
CC	autoimmune disease or accidental damage. The polypeptides and nucleotides
CC	may also be used in the diagnosis of the above conditions, and in drug
CC	screening techniques. The present sequence represents a novel human
CC	polypeptide of the invention.
XX	Sequence 156 AA;
SQ	Query Match 17.2%; Score 738; DB 22; Length 156;
	Best Local Similarity 100.0%; Pred. No. 3.6e-67;
	Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	89 RNVIYINETHRRHGWLARLSYVLFIOERDVHKGFMATNTENVLNSRRVQEAIAEVA 148
Db	12 RNVIYINETHRRHGWLARLSYVLFIOERDVHKGFMATNTENVLNSRRVQEAIAEVA 71



QY 149 ELNPGSQQQSKAVNKKKRILOEMVATVSPAMIRLTGWLKLFNSFFWNIQHK 208  
 |||||  
 Db 72 ELNPGSQQQSKAVNKKKRILOEMVATVSPAMIRLTGWLKLFNSFFWNIQHK 131  
 |||||  
 QY 209 GQLEWVKAATETNLPPLFLPVHRSH 233  
 |||||  
 Db 132 GQLEWVKAATETNLPPLFLPVHRSH 156  
 |||||

RESULT 12  
 AAY72134  
 ID AAY72134 standard; Protein; 827 AA.  
 XX  
 AC AAY72134;  
 XX  
 DT 24-APR-2001 (first entry)  
 XX  
 DE E. coli glycerol-3-phosphate acyltransferase with ER retention sequence.  
 XX  
 KW Triacylglyceride; TAG; glycerol-3-phosphate acyltransferase; GPAT;  
 KW fatty acid; oil; errs; endoplasmic reticulum retention sequence.  
 XX  
 OS Escherichia coli.  
 XX  
 PN WO200078974-A2.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 20-JUN-2000; 2000WO-CA00738.  
 XX  
 PR 21-JUN-1999; 99US-0139788.  
 PR 16-FEB-2000; 2000US-0182905.  
 XX  
 PA (CANA ) NAT RES COUNCIL CANADA.  
 XX  
 PI Jain RK, Mackenzie SL;  
 DR N-PSDB; AAD02333.  
 XX

Transforming organisms with DNA encoding glycerol-3-phosphate  
 acyltransferase to increase their natural oil and fat levels and alter  
 the fatty acid composition of their triacylglycerides -  
 Claim 20; Page 56-60; 60pp; English.  
 CC The patent discloses a method for increasing the triacylglyceride (TAG)  
 CC content of an organism and/or for modifying the fatty acid composition  
 CC of the TAG by expressing in the organism, a DNA encoding a protein with  
 CC glycerol-3-phosphate acyltransferase GPAT activity. This method is used  
 CC to increase the natural oil and fat levels and to alter the fatty acid  
 CC composition of TAGs in plants and yeast. It can be used to manipulate  
 CC oil synthesis in other organisms such as yeast, other fungus and algae  
 CC for producing commodity and specialty oils. Increasing the oil content  
 CC of feed quality grains reduces the need for adding exogenous fats in the  
 CC diets of animals and birds.  
 CC The present sequence is Escherichia coli glycerol-3-phosphate acyl-  
 CC transferase (GPAT) with endoplasmic reticulum retention sequence (errs).  
 CC The p18+errs gene encoding this sequence is used to modify the type of  
 CC fatty acid at the sn-1 position of triacylglycerides (TAG). This enables  
 CC the production of structured TAGs, in which the fatty acids occupying  
 CC each position may be controlled. It has implications for manipulating  
 CC the fat content in humans and other animals.  
 XX Sequence 827 AA;

Query Match 9.1%; Score 391.5; DB 22; Length 827;  
 Best Local Similarity 24.0%; Pred No. 3.1e-30;  
 Matches 147; Conservative 118; Mismatches 248; Indels 99; Gaps 21;  
 QY 76 DKFFNPSTPSGLRNVIVINETHRRHGWLARRLSVLFIQERDVHKGWFATNTVENVLN 135  
 |||||  
 |||||

Db 182 DSFVRFS-PSVSLRRMADEHGTDTKTTIAQKLA-RVARMHFARQLAAVGP-ELPARQDLFN 238  
 QY 136 SSRVQEAETAEVAALNPDGSAQOOSKAVN--KVKKAKRILOEMVATVSPAMIRLTGWL 193  
 |||||  
 Db 239 KLLASRAIAKAVED-----EARSKKLSHEKAQONAIALMEEETIANFSEYEMIRLTDRL 291  
 |||||  
 QY 194 LKLFNSFFWNIQHKQLEWVKAATETNLPPLFLPVHRSHDYLLLTFFILFCHNIKAPYI 253  
 |||||  
 Db 292 GFTWNRLYQGGINVHNA--ERVRLAHGDHGLVVPVCHSRHMDYLLLSVYVHOGVLVPPHI 349  
 |||||  
 QY 254 ASGNLNI-----PIFSTLIHKLGFFIRRRRLDEPDGRKDVLYRALLHGHIVELLRQOOF 309  
 |||||  
 Db 350 AAGINLNFWPAGPIF---RELGAFFIR---TFKGNK--LYSTVFFREYLGELFSRGYS 399  
 |||||  
 QY 310 LEIFLEGRSRSGKTSACAGLLSVVDLTSTNVIPDILIIIPVGISYDRIIE-GHYNGEQ 368  
 |||||  
 Db 400 VEYFVEGGRSPTGLLDPKGTLSMTQAMLRGGTRPTILPIYIGYEHVMEVGYAKEL 459  
 |||||  
 QY 369 LGKPKNESLWSVARGVIRMLRKNYGCVRVDPAPQFSLKEYLES-----OSQKPVSA 422  
 |||||  
 Db 460 RGATKEKESLPQMLRGLSKL--RNLGQGVNFGPMPMLMTYLNQHVDPDWRESIDPI--- 513  
 |||||  
 QY 423 SLEQALLPAILLPSRPDAADDEGRDTSINESRNATDESRLRLIANLAHILFTASKSCAI 482  
 |||||  
 Db 514 ---EAVRPWLTP-----TNNIAADLMVRINNAGAA 542  
 |||||  
 QY 483 MSHIVACLILLYRHRGIDLSLTVED---FFVMKEEVLARDFDLGFSGNSDEVVMAIQ 538  
 |||||  
 Db 543 NAMNLCTALLASQRSITREQLTEQLNCYLDLMRNVYSTDTVP-SASASELIDHALQ 601  
 |||||  
 QY 539 LLGNCVTITHTSRNDEFFITPTTVPVSFELNFSNGVLHVFIEMAITACSLYAVLNKRG 598  
 |||||  
 Db 602 M--NKFEVEKDTIGDIIILPREQAV---LMTYRNNAIAHMLVPLSLMAIV----- 647  
 |||||  
 QY 599 LGGTSTPPNLIISOELVRKAASLCYLLSNGTSLPCQTFYQVCHETVKGFIQVIGILT 658  
 |||||  
 Db 648 -----TQHRHISRDVLMHVNVLVPMKAEFLFLRWDRDELDPVIDALANEMQROGLITL 701  
 |||||  
 QY 659 AEHDDQEDISPS 670  
 |||||  
 Db 702 --QDELHINPA 711  
 |||||

RESULT 13  
 AAY72133  
 ID AAY72133 standard; Protein; 827 AA.  
 XX  
 AC AAY72133;  
 XX  
 DT 24-APR-2001 (first entry)  
 XX  
 DE Escherichia coli glycerol-3-phosphate acyltransferase (GPAT).  
 XX Triacylglyceride; TAG; glycerol-3-phosphate acyltransferase;  
 KW GPAT; fatty acid; oil.  
 XX Escherichia coli.  
 XX WO200078974-A2.  
 XX  
 XX 28-DEC-2000.  
 XX  
 XX 20-JUN-2000; 2000WO-CA00738.  
 XX  
 XX 21-JUN-1999; 99US-0139788.  
 XX 16-FEB-2000; 2000US-0182905.  
 XX  
 XX (CANA ) NAT RES COUNCIL CANADA.  
 XX Jain RK, Mackenzie SL;  
 XX WPI; 2001-041428/05.  
 XX N-PSDB; AAD02332.





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2004, 04:06:03 ; Search time 574.937 Seconds  
(without alignments)  
15003.329 Million cell updates/sec

Title: US-09-935-290-3  
Perfect score: 2484  
Sequence: 1 atggatgaatgcactgac.....ttctgagtttctgtgtgtg 2484

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2276164 seqs, 1736306516 residues  
Total number of hits satisfying chosen parameters: 4552328

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2484	100.0	2484	11	US-09-935-290-3
2	2484	100.0	3003	11	US-09-935-290-1
3	2484	100.0	4033	13	US-09-814-353-20793
4	2480.8	99.9	3210	13	US-10-288-252-26
5	1879.2	75.7	2646	10	US-09-917-800A-1648
6	581.6	23.4	612	10	US-09-833-381-1961
7	437	17.6	482	11	US-09-918-995-21644
8	385.4	15.5	399	11	US-09-918-995-7822
9	348.4	14.0	440	10	US-09-833-381-1964
10	346.6	14.0	396	10	US-09-833-381-1959
11	342.2	13.8	429	10	US-09-960-352-7384
12	337.6	13.6	407	10	US-09-960-352-8821
c 13	331.8	13.4	868	10	US-09-833-381-1958
c 14	302	12.2	418	10	US-09-960-352-742
c 15	268.6	10.8	379	13	US-09-814-353-16182

c 16	233	9.4	303	10	US-09-960-352-13086	Sequence 13086, A
c 17	209.6	8.4	264	10	US-09-960-352-6239	Sequence 6239, Ap
c 18	207.2	8.3	244	13	US-09-814-353-3482	Sequence 3482, Ap
c 19	207.2	8.3	244	13	US-09-814-353-9798	Sequence 9798, Ap
c 20	172.6	6.9	387	10	US-09-960-352-7634	Sequence 7634, Ap
c 21	129	5.2	308	10	US-09-833-381-1963	Sequence 1963, Ap
c 22	109	4.4	2682	13	US-10-288-252-27	Sequence 27, Appl
c 23	109	4.4	2755	13	US-10-288-252-27	Sequence 27, Appl
c 24	99.4	4.0	501	10	US-09-960-352-4061	Sequence 4061, Ap
c 25	75.2	3.0	467	13	US-10-027-632-297935	Sequence 297935, A
c 26	75.2	3.0	467	13	US-10-027-632-297936	Sequence 297936, A
c 27	75.2	3.0	467	14	US-10-027-632-297935	Sequence 297935, A
c 28	75.2	3.0	467	14	US-10-027-632-297936	Sequence 297936, A
c 29	75.2	3.0	634	13	US-10-027-632-41737	Sequence 41737, A
c 30	75.2	3.0	634	13	US-10-027-632-41738	Sequence 41738, A
c 31	75.2	3.0	634	14	US-10-027-632-41737	Sequence 41737, A
c 32	75.2	3.0	634	14	US-10-027-632-41738	Sequence 41738, A
c 33	62	2.5	442	10	US-09-833-381-1965	Sequence 1965, Ap
c 34	59.2	2.4	469	13	US-10-027-632-62086	Sequence 62086, A
c 35	59.2	2.4	469	13	US-10-027-632-62087	Sequence 62087, A
c 36	59.2	2.4	469	13	US-10-027-632-63435	Sequence 63435, A
c 37	59.2	2.4	469	13	US-10-027-632-63434	Sequence 63434, A
c 38	59.2	2.4	469	14	US-10-027-632-62086	Sequence 62086, A
c 39	59.2	2.4	469	14	US-10-027-632-62087	Sequence 62087, A
c 40	59.2	2.4	469	14	US-10-027-632-63434	Sequence 63434, A
c 41	59.2	2.4	469	14	US-10-027-632-63435	Sequence 63435, A
c 42	57.6	2.3	2262	12	US-10-369-493-33072	Sequence 33072, A
c 43	56.4	2.3	450	10	US-09-833-381-1960	Sequence 1960, Ap
c 44	56	2.3	2262	12	US-10-369-493-32927	Sequence 32927, A
c 45	56	2.3	2613	12	US-10-369-493-41238	Sequence 41238, A

ALIGNMENTS

RESULT 1  
US-09-935-290-3  
; Sequence 3, Application US/09935290  
; Publication No. US20030044948A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller-Libermann, Rosana  
; TITLE OF INVENTION: 56919, A NOVEL HUMAN ACYLTRANSFERASE AND USES THEREOF  
; FILE REFERENCE: MNI-186  
; CURRENT APPLICATION NUMBER: US/09/935,290  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/226,509  
; PRIOR FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2484  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-935-290-3

Query Match	100.0%	Score 2484;	DB 11;	Length 2484;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2484;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGATGAATCTGCACCTGACCCCTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCA	60	
Db	1	ATGGATGAATCTGCACCTGACCCCTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCA	60	
Qy	61	GAATACAGTGTGTGTCATGTAAACACACAAAGTGAAGTGGGTGAGTGGCTTTAGA	120	
Db	61	GAATACAGTGTGTGTCATGTAAACACACAAAGTGAAGTGGGTGAGTGGCTTTAGA	120	
Qy	121	CCACCGCTTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTTAATGAGTCGAAAGG	180	
Db	121	CCACCGCTTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTTAATGAGTCGAAAGG	180	
Qy	181	CCATTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGCTGGCAAAATTTTCAAC	240	

Db 181 CCAATTTGTTGGAAGATGTTGTTACTCTCTGCACTCCCGAGAGCTGGGCAAAATTTTCAAC 240  
Qy 241 CCAGATACCGCTCTTTGGGTTTTCGGGAATGTTATTTATATCAATGAAATCTCACACAAGA 300  
Db 241 CCGATATCCGCTCTTTGGGTTTTCGGGAATGTTATTTATATCAATGAAATCTCACACAAGA 300  
Qy 301 CACCGGATGGCTTTCGAAGACGCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTG 360  
Db 301 CACCGGATGGCTTTCGAAGACGCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTG 360  
Qy 361 CATAGGGCATGTTTGGCCACCAATGTGACGTAATAATGTGCTGAACAGACGATAGATACAA 420  
Db 361 CATAGGGCATGTTTGGCCACCAATGTGACGTAATAATGTGCTGAACAGACGATAGATACAA 420  
Qy 421 GAGCAATTCGAGAAGTGGCTGTGAATTAACCCCTGATGGTTCTGCCCGACGCAATCA 480  
Db 421 GAGCAATTCGAGAAGTGGCTGTGAATTAACCCCTGATGGTTCTGCCCGACGCAATCA 480  
Qy 481 AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTAAAGGATTTCTTCAAGAAATGGTTGCCACT 540  
Db 481 AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTAAAGGATTTCTTCAAGAAATGGTTGCCACT 540  
Qy 541 GTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGTAATAAATCTGTTCAACAGCTTC 600  
Db 541 GTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGTAATAAATCTGTTCAACAGCTTC 600  
Qy 601 TTTTGGAACTTCAAAATTCACAAAGTCAACTTGAAGATGGTTAAAGCTGCAACTGAGACG 660  
Db 601 TTTTGGAACTTCAAAATTCACAAAGTCAACTTGAAGATGGTTAAAGCTGCAACTGAGACG 660  
Qy 661 AATTTGGCGCTCTGTTTCTACCAAGTTTACAGATGCCATATTTGACTATCTGCTGCTCACT 720  
Db 661 AATTTGGCGCTCTGTTTCTACCAAGTTTACAGATGCCATATTTGACTATCTGCTGCTCACT 720  
Qy 721 TTCAATCTCTCTGCTCAATAACATCAAAAGCAACATACATGCTTCCAGGCAATAATCTCAAC 780  
Db 721 TTCAATCTCTCTGCTCAATAACATCAAAAGCAACATACATGCTTCCAGGCAATAATCTCAAC 780  
Qy 781 ATCCCAATCTCAGTACTGATCCATAAGCTTGGGGCTTCTTCAACGACGAAGCTC 840  
Db 781 ATCCCAATCTCAGTACTGATCCATAAGCTTGGGGCTTCTTCAACGACGAAGCTC 840  
Qy 841 GATGAAACACAGATGACGGAAGATGTTCTCTATAGAGCTTTGCTTCCATGGGCATATA 900  
Db 841 GATGAAACACAGATGACGGAAGATGTTCTCTATAGAGCTTTGCTTCCATGGGCATATA 900  
Qy 901 GTTGAATTAATCTTGACAGCAGCAATCTTGAGATCTTCCTGGAAGCAACGCTTCTAGG 960  
Db 901 GTTGAATTAATCTTGACAGCAGCAATCTTGAGATCTTCCTGGAAGCAACGCTTCTAGG 960  
Qy 961 AGTGGAAAACCTCTTGCTCGGGCAGGACTTTTGTTCAGTTGGGTAGATCTCTGTCT 1020  
Db 961 AGTGGAAAACCTCTTGCTCGGGCAGGACTTTTGTTCAGTTGGGTAGATCTCTGTCT 1020  
Qy 1021 ACCAATGTCTCCAGACATCTTGATATAACCTGTGGAAATCTCTATGATCGCATTATC 1080  
Db 1021 ACCAATGTCTCCAGACATCTTGATATAACCTGTGGAAATCTCTATGATCGCATTATC 1080  
Qy 1081 GAAGGTCACTACATGTGTGAACAACTGGGCAAACTTAAGAAGAAATGAGACCTGTGGAGT 1140  
Db 1081 GAAGGTCACTACATGTGTGAACAACTGGGCAAACTTAAGAAGAAATGAGACCTGTGGAGT 1140  
Qy 1141 GTAGCAGAGGTGTTATAGAAATGTTACGAAATAACTATGGTTGTGTCGAGTGGATTTT 1200  
Db 1141 GTAGCAGAGGTGTTATAGAAATGTTACGAAATAACTATGGTTGTGTCGAGTGGATTTT 1200  
Qy 1201 GCACAGCAATTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAAGAAACCGGTGTCTGCT 1260  
Db 1201 GCACAGCAATTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAAGAAACCGGTGTCTGCT 1260  
Qy 1261 CTACTTTCCCTGGAGCAAGCGTTGTTACAGCTATACCTTCAAGACCCAGTGATGCT 1320  
Db 1261 CTACTTTCCCTGGAGCAAGCGTTGTTACAGCTATACCTTCAAGACCCAGTGATGCT 1320

Qy 1321 GCTGATGAAGGTAGAGACACGTCCTCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db 1321 GCTGATGAAGGTAGAGACACGTCCTCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Qy 1381 CGAAGAGGTTGATTTGCAAAATCTGGCTGAGCATATTTCTATATTCATCTGCTAGCAAGTCTGT 1440  
Db 1381 CGAAGAGGTTGATTTGCAAAATCTGGCTGAGCATATTTCTATATTCATCTGCTAGCAAGTCTGT 1440  
Qy 1441 GCCATTATCTCCACACACATTTGGCTTGGCTTCCCTCTCTACAGACACAGGCGAGGAAAT 1500  
Db 1441 GCCATTATCTCCACACACATTTGGCTTGGCTTCCCTCTCTACAGACACAGGCGAGGAAAT 1500  
Qy 1501 GATCTCTCCACATTTGGTTCGAAGACTTCTTTGTGATGAAAGAGAAAGTCTCTGGCTCGTGAT 1560  
Db 1501 GATCTCTCCACATTTGGTTCGAAGACTTCTTTGTGATGAAAGAGAAAGTCTCTGGCTCGTGAT 1560  
Qy 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAGATGTAGTAATGCATGCCATACAGCTGCTG 1620  
Db 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAGATGTAGTAATGCATGCCATACAGCTGCTG 1620  
Qy 1621 GGAATTTGTTCAATCACACCTACCCACACTAGCAGGAACGATGATTTTATACCCCCAGC 1680  
Db 1621 GGAATTTGTTCAATCACACCTACCCACACTAGCAGGAACGATGATTTTATACCCCCAGC 1680  
Qy 1681 ACAACTGTCCCATCAGTCTTTCGAACTCAACTTCTACAGCAATGGGGTACTTCAATGCTTT 1740  
Db 1681 ACAACTGTCCCATCAGTCTTTCGAACTCAACTTCTACAGCAATGGGGTACTTCAATGCTTT 1740  
Qy 1741 ATCATGGAGGCCATCATAGCTTTGAGCCCTTTATGCAATGCTGAAAGAGGGGACTGGGG 1800  
Db 1741 ATCATGGAGGCCATCATAGCTTTGAGCCCTTTATGCAATGCTGAAAGAGGGGACTGGGG 1800  
Qy 1801 GGTCCCACTAGCACCCACCTAACCCTGATCAGCAGGACGAGCTGGTGGGAGGCGGCC 1860  
Db 1801 GGTCCCACTAGCACCCACCTAACCCTGATCAGCAGGACGAGCTGGTGGGAGGCGGCC 1860  
Qy 1861 AGCTGTGCTACTCTCTCCAATGAAGCACCATCTCACTGCTTGCCTTGCAGACATTTTAC 1920  
Db 1861 AGCTGTGCTACTCTCTCCAATGAAGCACCATCTCACTGCTTGCCTTGCAGACATTTTAC 1920  
Qy 1921 CAACTGCTGCATGAAACAGTAGAGAAAGTTTATCCAGATATGGCAATCTTACAGTGGCAGAG 1980  
Db 1921 CAACTGCTGCATGAAACAGTAGAGAAAGTTTATCCAGATATGGCAATCTTACAGTGGCAGAG 1980  
Qy 1981 CAGATGACAGGAAGATATCAGTCTAGTCTTGTGAGCAGCAGTGGGCAAGAGCTT 2040  
Db 1981 CAGATGACAGGAAGATATCAGTCTAGTCTTGTGAGCAGCAGTGGGCAAGAGCTT 2040  
Qy 2041 CCAGAACCTTTGCTTGGGAAAGTGAAGAGATGAAGAGATGAAAGATGAAAGATGAAAGATGAA 2100  
Db 2041 CCAGAACCTTTGCTTGGGAAAGTGAAGAGATGAAAGATGAAAGATGAAAGATGAAAGATGAA 2100  
Qy 2101 CAGCGAGATTGCTTACCTGAAGGTGAGCCCAATCCAAGGACCCAGCAGTTTATCACCTTC 2160  
Db 2101 CAGCGAGATTGCTTACCTGAAGGTGAGCCCAATCCAAGGACCCAGCAGTTTATCACCTTC 2160  
Qy 2161 TTACAGACACTCTCTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2220  
Db 2161 TTACAGACACTCTCTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2220  
Qy 2221 AACTTCAGTGGTCTCTCTCCAGAACTGATGATGAAAGATGTCAGAAATGTCAGAAATGTCAGAA 2280  
Db 2221 AACTTCAGTGGTCTCTCTCCAGAACTGATGATGAAAGATGTCAGAAATGTCAGAAATGTCAGAA 2280  
Qy 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATTTGCTTGTGAAG 2340  
Db 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATTTGCTTGTGAAG 2340  
Qy 2341 AATGCTGTGAAATGTTTAAAGGATATTTGGGTTTTCAGAGGACCCAAACAAAGAGAGTG 2400  
Db 2341 AATGCTGTGAAATGTTTAAAGGATATTTGGGTTTTCAGAGGACCCAAACAAAGAGAGTG 2400

Accession	Gene	Strain	Sequence	Position
QY	2401	TCTGTTT	TAGAACTGAGCGAGCACTTCTACCTCAATGCAACCGCAAAAACTTCTAGAA	2460
Db	2401	TCTGTTT	TAGAACTGAGCGAGCACTTCTACCTCAATGCAACCGCAAAAACTTCTAGAA	2460
QY	2461	TATATCT	GAGTTTTGTGGTGCTG	2484
Db	2461	TATATCT	GAGTTTTGTGGTGCTG	2484

## RESULT 2

```

US-09-935-290-1
; Sequence 1, Application US/09935290
; Publication No. US20030044948A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 56919, A NOVEL HUMAN ACYLTRANSFERASE AND USES THEREOF
; FILE REFERENCE: MNI-186
; CURRENT APPLICATION NUMBER: US/09/935,290
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/226,509
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3003
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (341)...(2827)
; NAME/KEY: misc_feature
; LOCATION: (1)...(3003)
; OTHER INFORMATION: n = A,T,C or G
US-09-935-290-1

```

Query Match	100.0%;	Score 2484;	DB 11;	Length 3003;				
Best Local Similarity	100.0%;	Prod. No. 0;						
Matches 2484;	Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;			
Qy	1	ATG	GAT	GAATCTGC	ACTGACCCCTGGTAC	AATAGATGTTCTTATCTGCCACATTTCATCA	60	
Db	341	ATG	ATGAATCTGC	ACTGACCCCTGGTAC	AATAGATGTTCTTATCTGCCACATTTCATCA	400		
Qy	61	GA	ATACAGTGT	TGGTTCGATGA	AGCACACAAGT	CAGGAATGGG	GTGAGTGTGGCTTTTGA	120
Db	401	GA	ATACAGTGT	TGGTTCGATGA	AGCACACAAGT	CAGGAATGGG	GTGAGTGTGGCTTTTGA	460
Qy	121	CCC	ACCGCTTTC	CAGATCTG	CAACTTTTAA	AATGGAAAGAAAGCCCTAA	TGAGTCGGAAAGG	180
Db	461	CCC	ACCGCTTTC	CAGATCTG	CAACTTTTAA	AATGGAAAGAAAGCCCTAA	TGAGTCGGAAAGG	520
Qy	181	CCA	TTTGTGGA	AGATGTTTACT	CTCTGCACTCCC	CAGAGCTGGG	ACAAATTTTTCAC	240
Db	521	CCA	TTTGTGGA	AGATGTTTACT	CTCTGCACTCCC	CAGAGCTGGG	ACAAATTTTTCAC	580
Qy	241	CCC	AGTATCCCGT	CTTTGGGTTTGG	GGAATGTTATATATCA	ATGAACTCACA	CAAGA	300
Db	581	CCC	AGTATCCCGT	CTTTGGGTTTGG	GGAATGTTATATATCA	ATGAACTCACA	CAAGA	640
Qy	301	CAC	CGCGATGGCT	TGCAAGACGCTTCTTAC	GTTCCTTTTATTC	CAAGACGAGATGTG	360	
Db	641	CAC	CGCGATGGCT	TGCAAGACGCTTCTTAC	GTTCCTTTTATTC	CAAGACGAGATGTG	700	
Qy	361	CAT	AAGGGCATGTTT	TGCCACCAATGTG	ACTGAAATGTGCTG	AACACGACGATGAGTACAA	420	
Db	701	CAT	AAGGGCATGTTT	TGCCACCAATGTG	ACTGAAATGTGCTG	AACACGACGATGAGTACAA	760	
Qy	421	GAG	GCAATTG	CAGAGTGGCTGC	TGCTGAATTA	AACCCCTGATGGTCTG	CTGCCACGACGACATCA	480
Db	761	GAG	GCAATTG	CAGAGTGGCTGC	TGCTGAATTA	AACCCCTGATGGTCTG	CTGCCACGACGACATCA	820
Qy	481	AA	AGCCGTTAA	CAAGTGA	AAAAAGAAAGCT	AAAAAGGATTTCTTCA	AGAAATTGGTTGCCACT	540

821	AAAGCCGTTTACAAAGTGAAAGAAAGCTAAAGAGATCTTTCAAGAAATGGTTGCCACT	880	
Qy	GTCTCACCGGCAATGATCAGACTGACCTGGGTGGGTGCTGCTATAAATCTGTTCAACAGCTTC	600	
Db	881	GTCTCACCGGCAATGATCAGACTGACCTGGGTGGGTGCTGCTATAAATCTGTTCAACAGCTTC	940
Qy	601	TTTTGGAAACATTTCAAATTTCAAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACG	660
Db	941	TTTTGGAAACATTTCAAATTTCAAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACG	1000
Qy	661	AAATTTGGCGGCTCTGTTTCTTACCAAGTTTCATAGATCCCATATTTGACTATCTGCTGCTCACT	720
Db	1001	AAATTTGGCGGCTCTGTTTCTTACCAAGTTTCATAGATCCCATATTTGACTATCTGCTGCTCACT	1060
Qy	721	TTCAATTTCTCTTCTGCCATAACATCAAAAGCACCATACATTTGCTTCAGGCATAAATCTCAAC	780
Db	1061	TTCAATTTCTCTTCTGCCATAACATCAAAAGCACCATACATTTGCTTCAGGCATAAATCTCAAC	1120
Qy	781	ATCCCAATTTCTCAGTACTCTTGATCCATTAAGCTTGGGGGCTTTCTTTCATACGACGAAGGCTC	840
Db	1121	ATCCCAATTTCTCAGTACTCTTGATCCATTAAGCTTGGGGGCTTTCTTTCATACGACGAAGGCTC	1180
Qy	841	GATGAACACCAAGATGAGCGGAAAGATGTTCTCTATAGAGCTTTGCTTCCATGGGCATATA	900
Db	1181	GATGAACACCAAGATGAGCGGAAAGATGTTCTCTATAGAGCTTTGCTTCCATGGGCATATA	1240
Qy	901	GTTGAAATTTACTTCGACAGCAGCAATCTTGAGATCTTCTGGAAGGCACACGTTTCTAGG	960
Db	1241	GTTGAAATTTACTTCGACAGCAGCAATCTTGAGATCTTCTGGAAGGCACACGTTTCTAGG	1300
Qy	961	AGTGGAAAAACCTCTGTGCTCGGCGCAGGACTTTTGTGAGTTGTGGTAGACTCTGTCT	1020
Db	1301	AGTGGAAAAACCTCTGTGCTCGGCGCAGGACTTTTGTGAGTTGTGGTAGACTCTGTCT	1360
Qy	1021	ACCAATGTCTATCCAGACATCTTGATTAATACCTGTTGGAATCTCTCTATGATCGCATATC	1080
Db	1361	ACCAATGTCTATCCAGACATCTTGATTAATACCTGTTGGAATCTCTCTATGATCGCATATC	1420
Qy	1081	GAAGGTCACTACAAATGTTGAACAACTCGGCAAACTTAAGAAGATGAGAGCTGTGGAGT	1140
Db	1421	GAAGTCACTTACAATGGTGAACAACTGGGCAAACTTAAGAAGATGAGAGCTGTGGAGT	1480
Qy	1141	GTAGCAAGAGGTGTTATTAGAAATGTTACGAAAAAACTATGTTGTGTCGAGTGGAATTTT	1200
Db	1481	GTAGCAAGAGGTGTTATTAGAAATGTTACGAAAAAACTATGTTGTGTCGAGTGGAATTTT	1540
Qy	1201	GCACAGCCATTTTCCTTTAAAGGAATATTAGAAAGCCAAAGTCAGAAAACCGGTGTCTGCT	1260
Db	1541	GCACAGCCATTTTCCTTTAAAGGAATATTAGAAAGCCAAAGTCAGAAAACCGGTGTCTGCT	1600
Qy	1261	CTACTTTCCCTGGAGCAAGCGTTGTTACAGCTATACTTCTTCAAGACCCAGTGATGCT	1320
Db	1601	CTACTTTCCCTGGAGCAAGCGTTGTTACAGCTATACTTCTTCAAGACCCAGTGATGCT	1660
Qy	1321	GCTGATGAAGGTAGAGACAGCTCTCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTTA	1380
Db	1661	GCTGATGAAGGTAGAGACAGCTCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTTA	1720
Qy	1381	CGAAGAGGTTGATTGCAAAATCTGGCTGAGCATATTCTATTCATCTGTAGCAAGTCTCTGT	1440
Db	1721	CGAAGAGGTTGATTGCAAAATCTGGCTGAGCATATTCTATTCATCTGTAGCAAGTCTCTGT	1780
Qy	1441	GCCATTTATGCCACACATTTGGCTTGCTGCTCTCTACAGACACAGCGAGGAAAT	1500
Db	1781	GCCATTTATGCCACACATTTGGCTTGCTGCTCTCTACAGACACAGCGAGGAAAT	1840
Qy	1501	GATCTCTCCCAATTTGGTTCGAAGACTTCTTTGTGATGAAAGAGGAAGTCTTGGCTCGTGAT	1560
Db	1841	GATCTCTCCCAATTTGGTTCGAAGACTTCTTTGTGATGAAAGAGGAAGTCTTGGCTCGTGAT	1900
Qy	1561	TTTGACCTGGGTTTCTCAGGAAATTCAGAAAGATGTAGTAATGCAATGCCATACAGCTGCTG	1620
Db	1901	TTTGACCTGGGTTTCTCAGGAAATTCAGAAAGATGTAGTAATGCAATGCCATACAGCTGCTG	1960



Qy	1621	GGAAATTGTTGTCACAATCACCCACACTAGCAGGAACGATGAGTTTTTATATCACCCCCAGC	1680
Db	1961	GGAAATTGTTGTCACATCACCCACACTAGCAGGAACGATGAGTTTTTATATCACCCCCAGC	2020
Qy	1681	ACAACTGTGCCATCAGTCTTCGAACCTCAACTCTACAGCAATGGGTACTTTCATGCTTTT	1740
Db	2021	ACAACTGTGCCATCAGTCTTCGAACCTCAACTCTACAGCAATGGGTACTTTCATGCTTTT	2080
Qy	1741	ATCATGGAGGCCATCATAGCTTCGAGCCTTTATGTCAGTTCTGAAACAAGAGGGGACTGGGG	1800
Db	2081	ATCATGGAGGCCATCATAGCTTCGAGCCTTTATGTCAGTTCTGAAACAAGAGGGGACTGGGG	2140
Qy	1801	GGTCCCACTAGACCCCACTTAACCTGATCAGCCAGGAGCAGCTGCTGCGGAAGGCGGCC	1860
Db	2141	GGTCCCACTAGACCCCACTTAACCTGATCAGCCAGGAGCAGCTGCTGCGGAAGGCGGCC	2200
Qy	1861	AGCCTGTGTACCTCTCTCCAATGAAGGCACCATCTCACTGCTCTGCCAGACATTTTAC	1920
Db	2201	AGCCTGTGTACCTCTCTCCAATGAAGGCACCATCTCACTGCTCTGCCAGACATTTTAC	2260
Qy	1921	CAAGTCTGCCATGAACAGTAGGAAAGTTTATCCAGTATGGCACTTTTACAGTGGCAGAG	1980
Db	2261	CAAGTCTGCCATGAACAGTAGGAAAGTTTATCCAGTATGGCACTTTTACAGTGGCAGAG	2320
Qy	1981	CAGATGAC CAGGAAGATATCAGTCTCTAGTCTTGTCAGCAGCAGTGGGACAAGAAGCTT	2040
Db	2321	CAGATGAC CAGGAAGATATCAGTCTCTAGTCTTGTCAGCAGCAGTGGGACAAGAAGCTT	2380
Qy	2041	CCAGAACCTTTGCTCTGGAGAAGTGATGAAGAAGATGAACACAGTGACTTTGGGGAGGAA	2100
Db	2381	CCAGAACCTTTGCTCTGGAGAAGTGATGAAGAAGATGAACACAGTGACTTTGGGGAGGAA	2440
Qy	2101	CAGCGAGATTGCTACCTGAAAGGTGAGCCAAATCCAAGGAGCACCAGCAGTTTATCACCTTC	2160
Db	2441	CAGCGAGATTGCTACCTGAAAGGTGAGCCAAATCCAAGGAGCACCAGCAGTTTATCACCTTC	2500
Qy	2161	TTACAGAGACTCCTTTGGGCCCTTTGCTGGAGGCCCTACAGCTCTGCTGCCATCTTTGTTTAC	2220
Db	2501	TTACAGAGACTCCTTTGGGCCCTTTGCTGGAGGCCCTACAGCTCTGCTGCCATCTTTGTTTAC	2560
Qy	2221	AACCTCAGTGGTCTGTTTCCAGAACCTTGAGTATCTGCAAAAGTTTGCACAAATACCTTAATA	2280
Db	2561	AACCTCAGTGGTCTGTTTCCAGAACCTTGAGTATCTGCAAAAGTTTGCACAAATACCTTAATA	2620
Qy	2281	ACCAGAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCCACATATTCCTTGTGAAG	2340
Db	2621	ACCAGAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCCACATATTCCTTGTGAAG	2680
Qy	2341	AATGCTGTGAAATGTTTAAAGGATATTTGGGGTTTTTCAAGGAGACCAACAAAAGAGAGTG	2400
Db	2681	AATGCTGTGAAATGTTTAAAGGATATTTGGGGTTTTTCAAGGAGACCAACAAAAGAGAGTG	2740
Qy	2401	TCTGTTTTTAGAACTGAGCAGCATTCTTCTACCTCAATGCAACCGACAAAAAATCTCTAGAA	2460
Db	2741	TCTGTTTTTAGAACTGAGCAGCATTCTTCTACCTCAATGCAACCGACAAAAAATCTCTAGAA	2800
Qy	2461	TATATCTGAGTTTTTGTGGTGCTG	2484
Db	2801	TATATCTGAGTTTTTGTGGTGCTG	2824

### RESULT 3

RESULT 3  
US-09-814-353-20793  
; Sequence 20793, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela  
; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER

Db 941 TTTTGGAACTTCAAAATTCACAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGAGC 1000  
 Qy 661 AATTTGCGCTTCTGTTTCTACAGTTCATAGATCCCATATTTAGTACTCTGCTGCTCACT 720  
 Db 1001 AATTTGCGCTTCTGTTTCTACAGTTCATAGATCCCATATTTAGTACTCTGCTGCTCACT 1060  
 Qy 721 TTTATTTCTTCTGTCATTAACATCAAGGACCATATGCTTTCAGGCAATTAATCTCAAC 780  
 Db 1061 TTTATTTCTTCTGTCATTAACATCAAGGACCATATGCTTTCAGGCAATTAATCTCAAC 1120  
 Qy 781 ATCCCAATCTTTCAGTACCTTGTATCATTAAGCTTTGGGGCTTCTTCATACGACGAGGCTC 840  
 Db 1121 ATCCCAATCTTTCAGTACCTTGTATCATTAAGCTTTGGGGCTTCTTCATACGACGAGGCTC 1180  
 Qy 841 GATGAACAACCCAGATGGAACGGAAGATGTTCTCTATAGAGCTTTTCTCCATGGGCATATA 900  
 Db 1181 GATGAACAACCCAGATGGAACGGAAGATGTTCTCTATAGAGCTTTTCTCCATGGGCATATA 1240  
 Qy 901 GTTGAATTACTTCGACAGCAGCAATTTCTTGAGATCTTCTGGAAGGACACAGTTCTAGG 960  
 Db 1241 GTTGAATTACTTCGACAGCAGCAATTTCTTGAGATCTTCTGGAAGGACACAGTTCTAGG 1300  
 Qy 961 AGTGAAGAAACCTCTTGTGCTCGGCAGGACTTTTGTCAAGTGTGGTAGACTCTGTCT 1020  
 Db 1301 AGTGAAGAAACCTCTTGTGCTCGGCAGGACTTTTGTCAAGTGTGGTAGACTCTGTCT 1360  
 Qy 1021 ACCAATGTATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATCGCAATATC 1080  
 Db 1361 ACCAATGTATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATCGCAATATC 1420  
 Qy 1081 GAAGTCTACTAATGTTGTAACACTGGGCAAACTTAAGAAATGAGAGCTGTGGAGT 1140  
 Db 1421 GAAGTCTACTAATGTTGTAACACTGGGCAAACTTAAGAAATGAGAGCTGTGGAGT 1480  
 Qy 1141 GTAGCAAGAGGTGTTATAGATGTTAGCAAAACTATGTTGTCCGAGTGAATTTT 1200  
 Db 1481 GTAGCAAGAGGTGTTATAGATGTTAGCAAAACTATGTTGTCCGAGTGAATTTT 1540  
 Qy 1201 GCACAGCCATTTTCTTAAAGGAATTTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCT 1260  
 Db 1541 GCACAGCCATTTTCTTAAAGGAATTTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCT 1600  
 Qy 1261 CTACTTTCCCTGGAGCAAGGTTGTTACAGCTATCTTCTTCAAGACCCAGTGAATGCT 1320  
 Db 1601 CTACTTTCCCTGGAGCAAGGTTGTTACAGCTATCTTCTTCAAGACCCAGTGAATGCT 1660  
 Qy 1321 GCTGATGAAGGTAGAGACAGTCCATTAATGATCCAGAAATGCAAGATGAATCCCTA 1380  
 Db 1661 GCTGATGAAGGTAGAGACAGTCCATTAATGATCCAGAAATGCAAGATGAATCCCTA 1720  
 Qy 1381 CGAAGGAGGTTGATTGCAAACTGCGCTGAGCATATTTCTATTCACTGCTAGCAAGTCTGT 1440  
 Db 1721 CGAAGGAGGTTGATTGCAAACTGCGCTGAGCATATTTCTATTCACTGCTAGCAAGTCTGT 1780  
 Qy 1441 GCCATTATGTCACACATTTGCTGCTTCCCTGCTCTCTACAGACACAGGAGGGAAT 1500  
 Db 1781 GCCATTATGTCACACATTTGCTGCTTCCCTGCTCTCTACAGACACAGGAGGGAAT 1840  
 Qy 1501 GATCTCTCCATTTGCTGAAGACTTTCTTGTGATGAAGAGGAAGTCTCGCTGCTGAT 1560  
 Db 1841 GATCTCTCCATTTGCTGAAGACTTTCTTGTGATGAAGAGGAAGTCTCGCTGCTGAT 1900  
 Qy 1561 TTTGACCTGGGTTCTCAGGAATTCAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1620  
 Db 1901 TTTGACCTGGGTTCTCAGGAATTCAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1960  
 Qy 1621 GGAATTTGTCACAAATCACCACTAGCAGGAACGATGATGATGATGATGATGATGATGATGATGAT 1680  
 Db 1961 GGAATTTGTCACAAATCACCACTAGCAGGAACGATGATGATGATGATGATGATGATGATGATGAT 2020  
 Qy 1681 ACACTGTCCTCATGCTTTCGAATCTTCAAGTCTTCAAGCAATGAGGATGATGATGATGATGAT 1740  
 Db 2021 ACACTGTCCTCATGCTTTCGAATCTTCAAGTCTTCAAGCAATGAGGATGATGATGATGATGAT 2080

Qy 1741 ATCATGGAGGCCATCATAGCTTTGAGCTTTTATGCAGTTCTGAAACAAGAGGGGACTGGGG 1800  
 Db 2081 ATCATGGAGGCCATCATAGCTTTGAGCTTTTATGCAGTTCTGAAACAAGAGGGGACTGGGG 2140  
 Qy 1801 GGTCCCACTAGCAGCCCACTTAAGCTGATGAGCAGGAGGAGCTGGTGGGAAAGCGGCGC 1860  
 Db 2141 GGTCCCACTAGCAGCCCACTTAAGCTGATGAGCAGGAGGAGCTGGTGGGAAAGCGGCGC 2200  
 Qy 1861 AGCTGTGCTACCTTCTTCCAAATGAAGGACCATCTCACTGCTTGCACACATTTTAC 1920  
 Db 2201 AGCTGTGCTACCTTCTTCCAAATGAAGGACCATCTCACTGCTTGCACACATTTTAC 2260  
 Qy 1921 CAAGTCTGCCATGAACAGTAGGAAAGTTTATCCAGTATGGCAATTTTACAGTGGCAGAG 1980  
 Db 2261 CAAGTCTGCCATGAACAGTAGGAAAGTTTATCCAGTATGGCAATTTTACAGTGGCAGAG 2320  
 Qy 1981 CACGATCACGAGGAAGATATCAGTCTTAGTCTTCTGAGCAGCAGTGGGACAAAGAGCTT 2040  
 Db 2321 CACGATCACGAGGAAGATATCAGTCTTAGTCTTCTGAGCAGCAGTGGGACAAAGAGCTT 2380  
 Qy 2041 CCAGAACTTTGCTTCTGAGAGAGTGAAGAGATGAAGACAGTGAATTTTGGGGAGGAA 2100  
 Db 2381 CCAGAACTTTGCTTCTGAGAGAGTGAAGAGATGAAGACAGTGAATTTTGGGGAGGAA 2440  
 Qy 2101 CAGCGAGATTGCTACCTGAAGGTGAGCCCAATCCAAAGGACCAAGCAGTATTTATCACCTTC 2160  
 Db 2441 CAGCGAGATTGCTACCTGAAGGTGAGCCCAATCCAAAGGACCAAGCAGTATTTATCACCTTC 2500  
 Qy 2161 TTACAGAGACTCTTGGGCTTTGCTGAGGCTTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 2220  
 Db 2501 TTACAGAGACTCTTGGGCTTTGCTGAGGCTTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 2560  
 Qy 2221 AACTTCAGTGGTCTGTTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAATA 2280  
 Db 2561 AACTTCAGTGGTCTGTTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAATA 2620  
 Qy 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATTTGTTGTGAAG 2340  
 Db 2621 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATTTGTTGTGAAG 2680  
 Qy 2341 AATCTGTGAAATGTTTAAAGGATATTTGGGTTTCAAGGAGACCAACAAAGAGAGTG 2400  
 Db 2681 AATCTGTGAAATGTTTAAAGGATATTTGGGTTTCAAGGAGACCAACAAAGAGAGTG 2740  
 Qy 2401 TCTGTTTGTAGACTGAGCAGCCTTTTCTACCTCAATGCAACCCGACAAAACCTTTCTAGAA 2460  
 Db 2741 TCTGTTTGTAGACTGAGCAGCCTTTTCTACCTCAATGCAACCCGACAAAACCTTTCTAGAA 2800  
 Qy 2461 TATATTCTGAGTTTGTGGTGCTG 2484  
 Db 2801 TATATTCTGAGTTTGTGGTGCTG 2824

RESULT 4  
 US-10-288-252-26  
 ; Sequence 26, Application US/10288252  
 ; Publication No. JS20030143686A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: INVITE GENOMICS, INC.  
 ; APPLICANT: IAL, Preeti G.  
 ; APPLICANT: TANG, Y. Tom  
 ; APPLICANT: YUE, Henry  
 ; APPLICANT: BURFORD, Neil  
 ; APPLICANT: GANDHI, Ameena R.  
 ; APPLICANT: WARREN, Bridget A.  
 ; APPLICANT: YAO, Monique G.  
 ; APPLICANT: TRIBOULEY, Catherine M.  
 ; APPLICANT: BAUGHN, Mariah R.  
 ; APPLICANT: LEE, Ernestine A.  
 ; APPLICANT: HAFALIA, April J.A.  
 ; APPLICANT: LU, Yan  
 ; APPLICANT: GRIFFIN, Jennifer A.

```
; APPLICANT: SANJANWALA, Madhu S.
; APPLICANT: DING, Li
; TITLE OF INVENTION: TRANSFERASES
; FILE REFERENCE: PI-0241 USA
; CURRENT APPLICATION NUMBER: US/10/288,252
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT US01/30424
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/252,819
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/249,639
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/247,931
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/246,001
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/244,025
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/238,481
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,523
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 26
; LENGTH: 3210
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030143686A1 1517230CB1
US-10-288-252-26

Query Match      99.9%; Score 2480.8; DB 13; Length 3210;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2482; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ATGATGAATCTGCACATGACCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 60
DB      338  ATGATGAATCTGCACATGACCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 397

QY      61  GAATACAGTGTGGTGCAGTAAAGACACACAAGTGAAGTGGGTGAGTGGCTTTAGA 120
DB      398  GAATACAGTGTGGTGCAGTAAAGACACACAAGTGAAGTGGGTGAGTGGCTTTAGA 457

QY      121  CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAATGAGTCGAAAGG 180
DB      458  CCCACCATCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAATGAGTCGAAAGG 517

QY      181  CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAATTTTCAAC 240
DB      518  CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAATTTTCAAC 577

QY      241  CCAGTATCCCGTCTTTGGGTTTGGCGAATGTTATTTATATCAATGAAATCCACAAAGA 300
DB      578  CCAGTATCCCGTCTTTGGGTTTGGCGAATGTTATTTATATCAATGAAATCCACAAAGA 637

QY      301  CACCGGATGGCTTGCAGACGCGCTTTCTTACGTTCTTTTATTAAGACGAGATGTG 360
DB      638  CACCGGATGGCTTGCAGACGCGCTTTCTTACGTTCTTTTATTAAGACGAGATGTG 697

QY      361  CATAAGGCGATGTTTGCACCAATGTGACCTGAAATGTGCTGAACAGCAGTACAGTACAA 420
DB      698  CATAAGGCGATGTTTGCACCAATGTGACCTGAAATGTGCTGAACAGCAGTACAGTACAA 757

QY      421  GAGGCAATTGAGAAGTGGCTGTGAATTAACCCCTGATGGTTCTGCCACAGCAATCA 480
DB      758  GAGGCAATTGAGAAGTGGCTGTGAATTAACCCCTGATGGTTCTGCCACAGCAATCA 817

QY      481  AAAGCCGTTAAAGAGTGAAGAAAGAGCTTAAAGGATTTCTTAAGAAATGGTTGCCACT 540
DB      818  AAAGCCGTTAAAGAGTGAAGAAAGAGCTTAAAGGATTTCTTAAGAAATGGTTGCCACT 877

541  GTCTCACCGCAATGATCAGACTGACTGGTGGGTGCTGCTAAAACTGTTCAACAGCTTC 600
DB      878  GTCTCACCGCAATGATCAGACTGACTGGTGGGTGCTGCTAAAACTGTTCAACAGCTTC 937

QY      601  TTTTGGAAACATTCAAATTCACAAAGGTCAACTTGAGATGGTTAAAGCTCAACTGAGCG 660
DB      938  TTTTGGAAACATTCAAATTCACAAAGGTCAACTTGAGATGGTTAAAGCTCAACTGAGCG 997

QY      661  AATTTGCCGCTTCTGTTTCTACAGTTCCATATCCCATATTTGACTATCTGCTGCTCACT 720
DB      998  AATTTGCCGCTTCTGTTTCTACAGTTCCATATCCCATATTTGACTATCTGCTGCTCACT 1057

QY      721  TTCAATTCTCTCTGCCATAAATCAAGACCAATACATTCATTGCTTCAGGCAATATCTCAAC 780
DB      1058  TTCAATTCTCTCTGCCATAAATCAAGACCAATACATTCATTGCTTCAGGCAATATCTCAAC 1117

QY      781  ATCCCAATCTTCAGTACCTTGATCCATTAAGCTTTGGGGGCTTCTTATACAGAGAGGCTC 840
DB      1118  ATCCCAATCTTCAGTACCTTGATCCATTAAGCTTTGGGGGCTTCTTATACAGAGAGGCTC 1177

QY      841  GATGAAACACACAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900
DB      1178  GATGAAACACACAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1237

QY      901  GTTGAATTACTTCGACAGCAGCAATTTCTTGGAGATCTTCTCGAAGGACACAGTTCTAGS 960
DB      1238  GTTGAATTACTTCGACAGCAGCAATTTCTTGGAGATCTTCTCGAAGGACACAGTTCTAGS 1297

QY      961  AGTGGAAAAACCTTGTGCTCGGCGAGGACTTTTGTTCAGTTGTGGTAGATCTCTGTCT 1020
DB      1298  AGTGGAAAAACCTTGTGCTCGGCGAGGACTTTTGTTCAGTTGTGGTAGATCTCTGTCT 1357

QY      1021  ACCAATGTCATCCAGACATCTTGATTAATACCTGTGGAATCTCTTATGATCGCATATC 1080
DB      1358  ACCAATGTCATCCAGACATCTTGATTAATACCTGTGGAATCTCTTATGATCGCATATC 1417

QY      1081  GAAAGTCACTACATAAGTGAACAACTGGGCAACCTTAAGAAAGTACAGAGCTGTGGAGT 1140
DB      1418  GAAAGTCACTACATAAGTGAACAACTGGGCAACCTTAAGAAAGTACAGAGCTGTGGAGT 1477

QY      1141  GTAGCAAGAGGTGTTATTAGAATGTTACGAAAAAACTATGGTTGTGTCGAGTGGATTTT 1200
DB      1478  GTAGCAAGAGGTGTTATTAGAATGTTACGAAAAAACTATGGTTGTGTCGAGTGGATTTT 1537

QY      1201  GCACAGCAATTTTCTTAAAGGAATATTAGAAAGCCAAAGTCAGAAACCGGTGCTGCT 1260
DB      1538  GCACAGCAATTTTCTTAAAGGAATATTAGAAAGCCAAAGTCAGAAACCGGTGCTGCT 1597

QY      1261  CTACTTTCTCCCTGGAGCAAGCTTTGTTACCCAGCTATATCTTCTTCAAGACCCAGTGTCT 1320
DB      1598  CTACTTTCTCCCTGGAGCAAGCTTTGTTACCCAGCTATATCTTCTTCAAGACCCAGTGTCT 1657

QY      1321  GCTGATGAAGGTAGAGACACGTCCTCAATTAATAGTCCAGAAATGCAACAGATGAATCCCTA 1380
DB      1658  GCTGATGAAGGTAGAGACACGTCCTCAATTAATAGTCCAGAAATGCAACAGATGAATCCCTA 1717

QY      1381  CGAAGAGGTGCTATGCAATCTGGCTGAGCATATTTCTTACTGCTAGCAAGTCTCTGT 1440
DB      1718  CGAAGAGGTGCTATGCAATCTGGCTGAGCATATTTCTTACTGCTAGCAAGTCTCTGT 1777

QY      1441  GCATTATGTCACACACATCTGCTTGGCTTGCCTGCTCTCTACAGACACAGGAGGAAAT 1500
DB      1778  GCATTATGTCACACACATCTGCTTGGCTTGCCTGCTCTCTACAGACACAGGAGGAAAT 1837

QY      1501  GATCTCTCCATTTGGTGCAGACCTTTTGTGTGATGAAGAGAAAGTCTCGCTCGTGTAT 1560
DB      1838  GATCTCTCCATTTGGTGCAGACCTTTTGTGTGATGAAGAGAAAGTCTCGCTCGTGTAT 1897

QY      1561  TTTGACCTGGGCTTCTCAGGAAATTCAGAGAGATGTAGTAATCATGCCATCAGCTGCTG 1620
DB      1898  TTTGACCTGGGCTTCTCAGGAAATTCAGAGAGATGTAGTAATCATGCCATCAGCTGCTG 1957

QY      1621  CGAAATTTGTGTCACAAATCACCCACACTAGCAGGAACGATGAGTGTGTTTTTATCACC 1680
```

```
Db 1958 GGAATTGTGTCACAAATCACCACACTAGCAGGAACGATGAGTTTTTATCACCCCGACG 2017
Qy 1681 ACAACTGTCCCATCAGTCTTGGAACCTCAACTTCTACAGCAATGGGGTACTTTCATGTCTTT 1740
Db 2018 ACAACTGTCCCATCAGTCTTGGAACCTCAACTTCTACAGCAATGGGGTACTTTCATGTCTTT 2077
Qy 1741 ATCATGGAGGCATCATAGCTTGCAGCCTTTATGCACTTCTGAAACAAGAGGGGACTGGG 1800
Db 2078 ATCATGGAGGCATCATAGCTTGCAGCCTTTATGCACTTCTGAAACAAGAGGGGACTGGG 2137
Qy 1801 GGTCCCACTAGCACCCCACTAACCTGATCAGCAGGAGCAGCTGGTGGGAGGCGGCC 1860
Db 2138 GGTCCCACTAGCACCCCACTAACCTGATCAGCAGGAGCAGCTGGTGGGAGGCGGCC 2197
Qy 1861 AGCTGTGCTACTTCTCTCCAATGAAGCGCACCATCTCACTGCCCTTGCAGACATTTTAC 1920
Db 2198 AGCTGTGCTACTTCTCTCCAATGAAGCGCACCATCTCACTGCCCTTGCAGACATTTTAC 2257
Qy 1921 CAACTGTCCCATGAACAAGTAGGAAAGTTTATCCAGTATGGGATTTCTTACAGTGGCAGAG 1980
Db 2258 CAACTGTCCCATGAACAAGTAGGAAAGTTTATCCAGTATGGGATTTCTTACAGTGGCAGAG 2317
Qy 1981 CACGATGACAGGAGATATCAGTCTAGTCTTGTGAGCAGCAGCTGGGGAACAAGAGCTT 2040
Db 2318 CACGATGACAGGAGATATCAGTCTAGTCTTGTGAGCAGCAGCTGGGGAACAAGAGCTT 2377
Qy 2041 CCAGAACCTTGTCTTGGAGAGTATGAAGAAGATGAAGACAGTGACTTTGGGGAGGAA 2100
Db 2378 CCTGAACCTTGTCTTGGAGAGTATGAAGAAGATGAAGACAGTGACTTTGGGGAGGAA 2437
Qy 2101 CAGCGAGATTCCTACCTCAAGGTGAGCCAAATCCAGGAGCACCAGCAGTTTATCACCTTC 2160
Db 2438 CAGCGAGATTCCTACCTCAAGGTGAGCCAAATCCAGGAGCACCAGCAGTTTATCACCTTC 2497
Qy 2161 TTACAGACATCCTTGGGCGCTTGTGAGGCGCTACAGCTCTGTGCCATCTTTGTTTAC 2220
Db 2498 TTACAGACATCCTTGGGCGCTTGTGAGGCGCTACAGCTCTGTGCCATCTTTGTTTAC 2557
Qy 2221 AACTTCAGTGTCTGTGTCAGAACCTCAGTATCTGCAAAAGTTGCAAAATACCTAATA 2280
Db 2558 AACTTCAGTGTCTGTGTCAGAACCTCAGTATCTGCAAAAGTTGCAAAATACCTAATA 2617
Qy 2281 ACCAGAACAGAAAGAAATTTGCAATATGTCAGAGTGCACATATTTGTTGTAAG 2340
Db 2618 ACCAGAACAGAAAGAAATTTGCAATATGTCAGAGTGCACATATTTGTTGTAAG 2677
Qy 2341 AATGCTGTGAAATGTTTAAGGATATTTGGGGTTTCAAGGAGACCAAAACAAGAGAGTG 2400
Db 2678 AATGCTGTGAAATGTTTAAGGATATTTGGGGTTTCAAGGAGACCAAAACAAGAGAGTG 2737
Qy 2401 TCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAATGCAACCGCACAAAATCTTCTAGAA 2460
Db 2738 TCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAATGCAACCGCACAAAATCTTCTAGAA 2797
Qy 2461 TATATCTGAGTTTGTGGTGTG 2484
Db 2798 TATATCTGAGTTTGTGGTGTG 2821
```

RESULT 5

US-09-917-800A-1648

; Sequence 1648, Application US/09917800A

; Patent No. US20020119462A1

; GENERAL INFORMATION:

; APPLICANT: Mendrick, Donna

; APPLICANT: Porter, Mark

; APPLICANT: Johnson, Kory

; APPLICANT: Castle, Arthur

; APPLICANT: Elashoff, Michael

; APPLICANT: Gene Logic, Inc.

; TITLE OF INVENTION: Molecular Toxicology Modeling

; FILE REFERENCE: 44921-5038-US

; CURRENT APPLICATION NUMBER: US/09/917,800A

; CURRENT FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: US 60/222,040

; PRIOR FILING DATE: 2000-07-31

; PRIOR APPLICATION NUMBER: US 60/222,880

; PRIOR FILING DATE: 2000-11-02

; PRIOR APPLICATION NUMBER: US 60/290,029

; PRIOR FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: US 60/290,645

; PRIOR FILING DATE: 2001-05-15

; PRIOR APPLICATION NUMBER: US 60/292,336

; PRIOR FILING DATE: 2001-05-22

; PRIOR APPLICATION NUMBER: US 60/295,798

; PRIOR FILING DATE: 2001-06-06

; PRIOR APPLICATION NUMBER: US 60/297,457

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/298,884

; PRIOR FILING DATE: 2001-06-19

; PRIOR APPLICATION NUMBER: US 60/303,459

; PRIOR FILING DATE: 2001-07-09

; NUMBER OF SEQ ID NOS: 1740

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1648

; LENGTH: 2646

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM\_017274

US-09-917-800A-1648

Query Match 75.7%; Score 1879.2; DB 10; Length 2646;  
Best Local Similarity 84.8%; Pred. No. 0;  
Matches 2106; Conservative 0; Mismatches 378; Indels 0; Gaps 0;

```
Qy 1 ATGATCAATCTGACATGACCCCTGGTACATAGATGTTCTTATCTGCCACATTCATCA 60
Db 160 ATGAGGAGTCTTCAGTGGACAAATGGGCACAATAGACGTTTCTTATCTGCCAATTCATCA 219
Qy 61 GAATACAGTGTGCTCCATGTAAAGCACACAAAGTGGGCTGAGTGGCTTTAGA 120
Db 220 GAATACAGCTTGGCCCATGTAAACACAGATGAGGACTGGGTGCTGGCTTTCAA 279
Qy 121 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAATAGTCGGAAGG 180
Db 280 CCTACCTCTTCAGATCCGCAACGCTGAATGGAAGAGAGCCCTCATGAGCGGAGAGG 339
Qy 181 CCATTTGTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAATTTTCAAC 240
Db 340 CCCTTCGTGGAGAGTGTGCTATTTCATGCACGCTCAGAGCTGGGAAAGGTTTTTCAAC 399
Qy 241 CCAGATATCCGCTCTTTGGGTTTGGGAAATGTTTATATCAATGAACTCACACAGA 300
Db 400 CCGAGTATCCATCTCTGGGTTTGGGAAATGTTTATATCAATGAACTCACACAGA 459
Qy 301 CACCGCGATGGCTTGAAGACGCTTTCTTACCTTTCTTTTATTCAGAGCGAGATGTG 360
Db 460 CACCGAGATGGCTGGCAAGACGCTTTCTTACATCTTTTGTTCAGAGCGGATGTC 519
Qy 361 CATAGGGCATGTTTGGCCCAATGTGCTGAAATGTGCTGAAACAGCAGTAGAGTCAA 420
Db 520 CACAAGGGCATGTTTGGCCCAATGTGCTGAAATGTGCTGAAACAGCAGTAGAGTCAA 579
Qy 421 GAGCAATTCAGAGTGGCTGTGAATTAACCCCTGATGTTCTGCTGCCAGCAATCA 480
Db 580 GAGCAATTCAGAGTGGCTGTGAATTAACCCCTGATGTTCTGCTGCCAGCAATCA 639
Qy 481 AAAGCCGTTAAACAAAGTGAAGAAAGAAAGGATTTCTTCAAGAAATGGTTGCCACT 540
Db 640 AAAGCCATCCAGAAAGTGAAGAAAGAAAGGAGATCTCCAGGAAATGGTTGCTACA 699
Qy 541 GTCTCACCGGCAATGATCAGACTGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 700 GTCTCCCCGGGATGATCAGCTGACTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 759
```

Qy	601	TTTTGGAA	CAATTC	AAATTC	CAAAATTC	CAAAAGGTC	AACTTG	GAGATG	GTTTAA	AGCTG	CAACTG	GAGACG	660
Db	760	TTCTGG	AACATTC	CAGATT	CAACAG	GGGTCA	ACTTG	GAGATG	TGTTGA	AGCTG	CAACTG	GAGACG	819
Qy	661	AAATTTG	CGCGTCT	TGTTTCT	TACCA	GTTC	ATAGAT	TCCCAT	ATTG	AGCTAT	CTGCT	CTCACT	720
Db	820	AATCTG	CGCGTCT	TGTTTCT	TGCGCG	TGC	ACAGAT	TCCC	CACTCG	AGTACT	TGCT	CTCACC	879
Qy	721	TTCAAT	CTCTTCT	GCCATA	CAATCA	AAAGCAC	CAATAT	GTCTT	CAGG	CAATAA	TCTCT	CAAC	780
Db	880	TTCAAT	CTCTTCT	GCCACA	CAATCA	AAAGCT	CCATAT	CGCCT	CGG	CACACAC	CTCT	CAAC	939
Qy	781	ATCCAA	TCTT	CAGTAC	CTTG	ATCC	ATAAGCT	TGGG	GGCTTCT	TTCAT	ACG	AGAGCTC	840
Db	940	ATCCCA	TCTT	CAGTAC	CTTG	ATTC	AAAGCT	TTC	CAAA	GGCTTTC	AT	TAAG	999
Qy	841	GATGA	AAAC	CCAGAT	TGAC	CGGA	AGATGTTCT	CTAT	AGAGCTT	TGCTT	CCAT	TGGGCAT	900
Db	1000	GACGAA	AACTC	CAGATG	GACG	CAAA	AGACATTTCT	GTAT	CAGAGCG	TGTGCT	CCAT	TGGGCAT	1059
Qy	901	GTTGAA	TTACTT	CGAC	AGCAG	CAATCTT	TGAGAT	CTTCT	TGGA	AGGCAC	ACAGT	TCTTAGG	960
Db	1060	GTTGAA	CTCCTC	CGAC	AGCAG	CAGTCTT	CGAGAT	CTTCT	TGGA	AGGCAC	CCCGCT	CCCGC	1119
Qy	961	AGTGG	AAAAA	ACCTT	TGTG	CTCGG	CAGBAC	TTTT	TGTCA	GTTG	TGG	TAGATAC	1020
Db	1120	AGTGG	CAAGAC	CTCCT	TG	CCCGG	CGGGCT	CTG	TCA	GTTG	TGG	TAGTCT	1179
Qy	1021	ACCAAT	GTCA	TCC	CAGAC	ATCTT	GTAT	AAATAC	CTGTT	TGGA	ATCTC	CTAT	1080
Db	1180	TCCAA	CACCAT	CCCT	GCAT	CTCCT	TGG	GCATCT	CGT	TATGAT	TCG	ATATC	1239
Qy	1081	GAAG	GTCA	TAC	AT	TG	TGTA	CAAC	AACTT	AGAA	AGAA	GTAG	1140
Db	1240	GAAG	GTCA	TAC	AT	TG	TGTA	CAAC	AGTCT	TGG	CAAG	CCCA	1299
Qy	1141	GTAG	CAAG	AGG	GT	TAT	TAGAA	TGTT	AC	GA	AAAA	AACTAT	1200
Db	1300	GTGCA	AGAG	CGGT	TAT	CAGAA	TCTC	CGG	AAAA	AACTAT	CGG	CTAT	1359
Qy	1201	GCAC	AG	CCATTT	CTT	TA	AAAG	GAATAT	T	TAG	AA	GC	1260
Db	1360	GCAC	AG	CCATTT	CTT	TG	AA	AGGA	TAT	T	TAG	AA	1419
Qy	1261	CTACT	TTTCC	TGG	AG	CAAG	CGT	TGTT	TAC	AGCTAT	TACT	CTT	1320
Db	1420	CCC	CTCT	CTTT	TGG	AG	CAAG	CAC	TGTT	TAC	AGCAAT	CTT	1479
Qy	1321	GCTG	ATGA	AGG	TAG	AGAC	ACG	TCT	CA	TAT	TAG	AGT	1380
Db	1480	GCTG	CCG	AA	CA	TGA	AGAC	CA	TG	TC	CA	TAT	1539
Qy	1381	CGA	AGG	AGG	T	GAT	TG	CAAA	TCT	GG	CTG	AG	1440
Db	1540	CGA	AGG	AGG	T	GAT	TG	CAAA	TCT	GG	CTG	AG	1599
Qy	1441	GCCA	TTAT	TG	CC	AC	AC	CA	CA	TAT	TG	TGG	1500
Db	1600	GCTAT	CAT	G	TCC	AC	CC	CA	CA	TAT	TG	TGG	1659
Qy	1501	GAT	CTCT	CC	CA	TAT	TG	TGG	TG	CGA	AG	ACT	1560
Db	1660	CAC	CTCT	CC	CA	TAT	TG	TGG	TG	CGA	AG	ACT	1719
Qy	1561	TTTGA	CC	TGG	GT	TCT	CAG	GA	AA	TTC	AG	AA	1620
Db	1720	TTTGA	CC	TGG	GT	TCT	CAG	GA	AA	TTC	AG	AA	1779
Qy	1621	GGAA	TTGT	GT	CA	AA	TTC	AG	AA	TTC	AG	AA	1680
Db	1780	GGAA	TTGT	GT	CA	AA	TTC	AG	AA	TTC	AG	AA	1839

Qy	1681	ACAACTGTCCCATCAGTCTTCGAACTCAACTTCTACAGCAATGGGGTACTTTCATGTCTTT	174
Db	1840	ACAACTGTCCCGTCCGCTTTTGAACCTTCTACAGCAATGGGGTACTTTCATGTCTTT	1899
Qy	1741	ATCATGAGGCCATCATAGCTTCGAGCCTTTATGTCAGTTCCTGAAACAGAGGGGACTGGG	1800
Db	1900	ATCATGGAAGCCATCATAGCTTCGAGCATTTATGTCAGTTCAGAAATAGAGGGGTTCCGGA	1959
Qy	1801	GGTCCCACTAGCACCCACCTAACTTGATCAGCCAGGAGCAGCTGGTGCAGAAAGGCGGCC	1860
Db	1960	GGGTCTCGCGGAGGCCCTTGGCAACCTGATCAGCCAGGAGCAGCTGGTGCAGAAAGGCGCC	2019
Qy	1861	AGCCTGTGCTACCTTCTCTCCATGAAGGCACCATCTCACTGCCTTGCAGACATTTTAC	1920
Db	2020	AGCCTGTGCTACCTTCTCTCTAAATGAAGGTACCAATTTCTCTGCCCCGCCAGACATTTTAC	2079
Qy	1921	CAAGTCTGCCATGAACAGTAGGAAAGCTTTTATCCAGTATGGCATTTCTACAGTGGCAGAG	1980
Db	2080	CAGGTTGTCTCAAGAGACAGTAGGAAAGTTTCATCCAGTACGGAAATTTCTCAGTGGCAGAG	2139
Qy	1981	CAGCATGACAGGAAGATATCAGTCTCTAGTCTTGTCTGAGCAGCAGTGGGACAGAAAGCTT	2040
Db	2140	CAAGATCACCAGGAAGATGTCACTCTCTGGCCTTCTCAGAGCAGCAGTGGAAACAAGAACTT	2199
Qy	2041	CCAGAACCTTTGTCTTGAGCAAGTGATGAAGAAAGATGAACACAGTGAATTTTGGGAGGAA	2100
Db	2200	CCGAGGCTCTGAACTGGAGAGTGACGAAGAAGATGAGGACAGTGAATTTTGTGAGGAG	2259
Qy	2101	CAGCGAGATTGTCTACCTGAAGAGGTGAGCCAAATCCAAGGAGCACGACGAGTTTATCACCTTC	2160
Db	2260	CAGCGTGAATTGCTACCTGAAAGGTGAGCCAGGCCAAAGGAGCACGACGAAATTCATCACCTTT	2319
Qy	2161	TTACAGAGACTCCTTTGGGCCCTTTGTGAGGCGCTACAGCTCTGCTGCCATCTTTGTTTAC	2220
Db	2320	CTGCAGAGGCTTCTGGGGGCCCTTGCTAGAAAGCCCTACAGCTCTGCTGCTCATCTTTGTCCAC	2379
Qy	2221	AACCTTCAGTGGTCTGTCTTCAGAACTGAGTATCTGCAAAAGTTGCAAAATACCTAAATA	2280
Db	2380	ACCTTCGCGGCCAGTCCCGGAGTCTGAGTACTCTGCAAGAAGTGCACAGGTACCTTCTC	2439
Qy	2281	ACCAGAAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCCACATATGTCTTGTGAAG	2340
Db	2440	ACCAGAGCGAGAGGAACGTGCGGTGTACGCTGAGAGTGCCACATCTGTCTTGTGAAG	2499
Qy	2341	AATGCTGTGAAATGTTTAAAGTATTTGGGGTTTTTCAAGGAGCCAAACAAAGAGAGTG	2400
Db	2500	AATGCTGTGAAATGTTTAAAGTATCTCGGGTTTTTCAAGAGACCAACAGCAGAAAGCGCG	2559
Qy	2401	TCTGTTTTTGAACCTGACAGCACTTTTCTTACTCTCAATGTCAACCGACCAAAATCTTCTAGAA	2460
Db	2560	TCGTCTTAGAATGAGCACCATTTTCTTACTCTCAGGCGAGCCGGCAGAGGCTCTCTGAA	2619
Qy	2461	TATATTCGTGATTTTGTGGTGCTG	2484
Db	2620	TACATTCGAGCTTCGTGGTGCTG	2643

## RESULT 6

```

US-09-833-381-1961
; Sequence 1961, Application US/09833381
; Patent No. US2020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US2020132090A1el
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1961
; LENGTH: 612

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1961

Query Match      23.4%; Score 581.6; DB 10; Length 612;
Best Local Similarity 99.0%; Pred. No. 9.1e-175;
Matches 606; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy 1061 TCTCCTATGATCGCATTCGAAGTCTACTCAATGTGTAACAACTGGGCAAACTTAAGA 1120
Db 1 TCCCGTATGATCGCATTCGAAGTCTACTCAATGTGTAACAACTGGGCAAACTTAAGA 60

Qy 1121 AGAATGAGACCTGTGAGTGTAGCAAGAGTGTATTAGAAATGTTACGAAAAAATATG 1180
Db 61 AGAATGAGACCTGTGAGTGTAGCAAGAGTGTATTAGAAATGTTACGAAAAAATATG 120

Qy 1181 GTTGTGTCGAGTGTGATTTTGCACAGCCATTTTCTTAAAGGAATATTAGAAAGCCAAA 1240
Db 121 GTTGTGTCGAGTGTGATTTTGCACAGCCATTTTCTTAAAGGAATATTAGAAAGCCAAA 180

Qy 1241 GTCAGAAA - CCGGTGTCTGCTCTACTTTCCCTGGAGCAAGCGTTGTACCAAGCTATACTT 1299
Db 181 GTCAGAAAACCGGTGTCTGCTCTACTTTCCCTGGAGCAAGCGTTGTACCAAGCTATACTT 240

Qy 1300 CCTTCAAGACCCAGTGTGCTGTGATGAAGGTAGAGACACGTCCATTATGAGTCCAGA 1359
Db 241 CCTTCAAGACCCAGTGTGCTGTGATGAAGGTAGAGACACGTCCATTATGAGTCCAGA 300

Qy 1360 AATCAACAGATGAATCCCTACGAGGAGTTGATGCAAACTCGGCTGAGCATATTCTA 1419
Db 301 AATCAACAGATGAATCCCTACGAGGAGTTGATGCAAACTCGGCTGAGCATATTCTA 360

Qy 1420 TTCACTGCTAGCAAGTCTGTGCCATTATGTCCACACACATTTGTGGCTTGCCTCTCTC 1479
Db 361 TTCACTGCTAGCAAGTCTGTGCCATTATGTCCACACACATTTGTGGCTTGCCTCTCTC 420

Qy 1480 TACAGACAGCGAGGGAATGATCTCTCCACATTGTGTCGAAGACTTCTTTTGTGATGAAA 1539
Db 421 TACAGACAGCGAGGGAATGATCTCTCCACATTGTGTCGAAGACTTCTTTTGTGATGAAA 480

Qy 1540 GAGGAAGTCTCGCTCGTGAATTTTACCTGGGTCTCAGGAATTCAGAGATGT-AGT 1598
Db 481 GGGGAAGTCTCGCTCGTGAATTTTACCTGGGTCTCAGGAATTCAGAGATGT-AGT 540

Qy 1599 AATCATGCCATACAGCTGTCTGGGAAATTTGTCAATACCCACACTAGCAGGAACGA 1658
Db 541 AATCATGCCATACAGCTGTCTGGGAAATTTGTCAATACCCACACTAGCAGGAACGA 600

Qy 1659 TGAGTTTTTTAT 1670
Db 601 TGAGTTTTTTAT 612
```

```

RESULT 7
US-09-918-995-21644
; Sequence 21644, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21644
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

Query Match      15.5%; Score 385.4; DB 11; Length 399;
Best Local Similarity 99.5%; Pred. No. 4.5e-112;
Matches 397; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 290 CTCACACAAGACACCGGGATGGCTTGCAGAGCCCTTTCTTACGTTCTTTTATCAAG 349
Db 1 CTCACACAAGACACCGGGATGGCTTGCAGAGCCCTTTCTTACGTTCTTTTATCAAG 60

Qy 350 AGCGAGATGTGCATTAAGGGGATGTTTCCACCACCAATGTGACTGAAAAATGTGC-TGAACAGC 408
Db 1 AGCGAGATGTGCATTAAGGGGATGTTTCCACCACCAATGTGACTGAAAAATGTGC-TGAACAGC 408
```

```

; NAME/KEY: misc_feature
; LOCATION: (1)...(482)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-21644

Query Match      17.6%; Score 437; DB 11; Length 482;
Best Local Similarity 100.0%; Pred. No. 1.6e-128;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 264 GCGGAATGTTATTATATCAATGAAACTCACACAAGACACCGCGGATGGCTTGCAGAGC 323
Db 46 GCGGAATGTTATTATATCAATGAAACTCACACAAGACACCGCGGATGGCTTGCAGAGC 105

Qy 324 CCTTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTGCATAAAGGCGCATGTTTGCACCAA 383
Db 106 CCTTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTGCATAAAGGCGCATGTTTGCACCAA 165

Qy 384 TGTGACTGAAATATGCTGTAACAGCAGTAGAGTCAAGAGCAATTTGAGAAAGTGGCTGC 443
Db 166 TGTGACTGAAATATGCTGTAACAGCAGTAGAGTCAAGAGCAATTTGAGAAAGTGGCTGC 225

Qy 444 TGAATTAACCCCTGATGGTTCTGCCAGCAGCAATCAAAAGCCGTTAAACAAGTGAATAA 503
Db 226 TGAATTAACCCCTGATGGTTCTGCCAGCAGCAATCAAAAGCCGTTAAACAAGTGAATAA 285

Qy 504 GAAAGCTAAAAGGATTTCTTCAAGAAATGGTTGCCACTGTCTCACCGGCAATGATCAGACT 563
Db 286 GAAAGCTAAAAGGATTTCTTCAAGAAATGGTTGCCACTGTCTCACCGGCAATGATCAGACT 345

Qy 564 GACTGGTGGTGTCTGCTAAAACCTGTTCAACAGCTTTCTTTTGGAAACATTTCAAAATTCACAA 623
Db 346 GACTGGTGGTGTCTGCTAAAACCTGTTCAACAGCTTTCTTTTGGAAACATTTCAAAATTCACAA 405

Qy 624 AGGTCAACTTGTAGATGTTAAAGCTCAACTGAGACCAATTTGCGGCTTCTGTTTCTACC 683
Db 406 AGGTCAACTTGTAGATGTTAAAGCTCAACTGAGACCAATTTGCGGCTTCTGTTTCTACC 465

Qy 684 AGTTCATAGATCCCATTA 700
Db 466 AGTTCATAGATCCCATTA 482

RESULT 8
US-09-918-995-7822
; Sequence 7822, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7822
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-7822

Query Match      15.5%; Score 385.4; DB 11; Length 399;
Best Local Similarity 99.5%; Pred. No. 4.5e-112;
Matches 397; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 290 CTCACACAAGACACCGGGATGGCTTGCAGAGCCCTTTCTTACGTTCTTTTATCAAG 349
Db 1 CTCACACAAGACACCGGGATGGCTTGCAGAGCCCTTTCTTACGTTCTTTTATCAAG 60

Qy 350 AGCGAGATGTGCATTAAGGGGATGTTTCCACCACCAATGTGACTGAAAAATGTGC-TGAACAGC 408
Db 1 AGCGAGATGTGCATTAAGGGGATGTTTCCACCACCAATGTGACTGAAAAATGTGC-TGAACAGC 408
```



Db 61 AGCAGATGTGCATAGAGGCGATGTTTCCACCAATGTGACTGAAATGTGCTTGAACAGC 120  
QY 409 AGTAGAGTACAGAGGCAATTGAGAGTGGCTGCTGAATTAACCCCTGATGTTCTGCC 468  
Db 121 AGTAGAGTACAGAGGCAATTGAGAGTGGCTGCTGAATTAACCCCTGATGTTCTGCC 180  
QY 469 CAGCAGCAATCAAAAGCCGTTTACAAAGTGAAAAGAAAGCTAAAGAGGATTTCTTCAAGAA 528  
Db 181 CACCAGCAATCAAAAGCCGTTTACAAAGTGAAAAGAAAGCTAAAGAGGATTTCTTCAAGAA 240  
QY 529 ATGGTTGCCACTGCTCTCACCGCAATGATCAGACTGACTGGGTGGTGTCTGCTAAAGCTG 588  
Db 241 ATGGTTGCCACTGCTCTCACCGCAATGATCAGACTGACTGGGTGGTGTCTGCTAAAGCTG 300  
QY 589 TTCAACAGCTCTTTTGGACATTCATCAAAAGTCAAAAGGTCACCTTCCAGATGTTAAAGCT 648  
Db 301 TTCAACAGCTCTTTTGGACATTCATCAAAAGTCAAAAGGTCACCTTCCAGATGTTAAAGCT 360  
QY 649 GCAACTGAGACGAATTTGCCGCTTCTGTTTCTACCAAGTT 687  
Db 361 GCAACTGAGACGAATTTGCCGCTTCTGTTTCTACCAAGTT 399  
RESULT 9  
US-09-833-381-1964  
; Sequence 1964, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1964  
; LENGTH: 440  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(440)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1964  
Query Match 14.0%; Score 348.4; DB 10; Length 440;  
Best Local Similarity 98.3%; Pred. No. 3.5e-100;  
Matches 349; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 87 CACAAGTAGGAATGGGGTGAAGTGGCTTTAGACCCACCGTCTTCAGATCTGCAACTTT 146  
Db 15 CCCAAGTGAGAAATGGGGTGAAGTGGCTTTAGACCCACCGTCTTCAGATCTGCAACTTT 74  
QY 147 ABAATGGAAGAAAGCCTATAGTTCGGAAGGCCATTTGTTGGAAGATGTTGTACTC 206  
Db 75 ABAATGGAAGAAAGCCTATAGTTCGGAAGGCCATTTGTTGGAAGATGTTGTACTC 134  
QY 207 CTGCACCTCCCGAGAGTGGGACAAAATTTTCAACCCCGATATCCCGTCTTTGGGTTGGC 266  
Db 135 CTGCACCTCCCGAGAGTGGGACAAAATTTTCAACCCCGATATCCCGTCTTTGGGTTGGC 194  
QY 267 GAATGTTATTTATCATATGAATCAACAGACACCGCGGATGGCTTGCAGAGCGCT 326  
Db 195 GAATGTTATTTATCATATGAATCAACAGACACCGCGGATGGCTTGCAGAGCGCT 254  
QY 327 TTCTTAGGTTCTTTTATTCAGAGCGAGATGTCATAGGCGATGTTGCCACCAATGT 386  
Db 255 TTCTTAGGTTCTTTTATTCAGAGCGAGATGTCATAGGCGATGTTGCCACCAATGT 314  
QY 387 GACTGAAATGTGCTGAACAGCAGTAGAGTACAGAGGCAATTCAGAAAGTGGCT 441

Db 315 GACTGAAATGTGCTTAAACNGCANTAGAGTACAGAGGCAATTCGNNAAAGTGGCT 369  
RESULT 10  
US-09-833-381-1959  
; Sequence 1959, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1959  
; LENGTH: 396  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(396)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1959  
Query Match 14.0%; Score 346.6; DB 10; Length 396;  
Best Local Similarity 95.9%; Pred. No. 1.2e-99;  
Matches 375; Conservative 0; Mismatches 13; Indels 3; Gaps 2;  
QY 1838 AGCAGCTGGTGGGAAAGCGCGCCAGCTGTGTACCTTCTCTCAATGAAGCACCATCT 1897  
Db 1 ACACAGCTGGTGGGAAAGCGCGCCAGCTGTGTACCTTCTCTCAATGAAGCACCATCT 60  
QY 1898 CACTGCTTGGCCAGACATTTTACCAAGTCTGCCATGAACAGTAGGAAAGTTATCCAGT 1957  
Db 61 CACTGCTTGGCCAGACATTTTACCAAGTCTGCCATGAACAGTAGGAAAGTTATCCAGT 120  
QY 1958 ATGGCATTTTACAGTGGCAGACGATGACCCAGAGATATCAGTCTTCTGCTG 2017  
Db 121 ATGGCATTTTACAGTGGCAGACGATGACCCAGAGATATCAGTCTTCTGCTG 180  
QY 2018 AGCAGCTGGGACAAAGAGCTTCCAGAACCTTTCTTGGAGAGTGTATGAAGAGATG 2077  
Db 181 AGCAGCTGGGACAAAGAGCTTCTTGAACCTTTCTTGGAGAGTGTATGAAGAGATG 240  
QY 2078 AAGACAGTGAATTTGGGAGGAAACAGCGAGATTGCTACCTGAA--GGTGAGCAATCCAA 2135  
Db 241 AAGACAGTGAATTTGGGAGGAAACAGCGAGATTGCTACCTGAAAGGTGAGCCCAATCCAA 300  
QY 2136 GGAGCACCAGCAGTTTATCACCCTTTTACAGAGACTCTTGGGCTTTGCT--GGAGGCT 2194  
Db 301 GGAGCACCAGCAGTTTAAACAACCTTCTTACAGAGACTCTTGGGCTTTGCTGGAGGCT 360  
QY 2195 ACAGCTCTGCTGCCATCTTTGTTTCACACTT 2225  
Db 361 ACAGCTCTGCTGCCATCTTTGTTTCACACTT 391  
RESULT 11  
US-09-960-352-7384  
; Sequence 7384, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Ningbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352



```
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7384
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 32-LIB2809-019-Q1-E1-H11
US-09-960-352-7384

Query Match      13.8%; Score 342.2; DB 10; Length 429;
Best Local Similarity 90.2%; Pred. No. 3.3e-98;
Matches 378; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

Qy 1330 GGTAGAGACAGTCCATTAAAGAGTCCAGAAATGCAACAGATGAATCCCTACGAGGAGG 1389
Db 14 GATACAGACAAGTCCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCGCGCA---AGA 70

Qy 1390 TTGATTGCAATTCGGCTGAGCATATCTATTCACTGCTAGCAAGTCTGTGCCATTATG 1449
Db 71 CTGATTGCAATTCGGCTGAGCATATCTTTCTACTGTAGCAAGTCTGTGCTATTATG 130

Qy 1450 TCCACACATTCGGCTTGGCTGCTCTCTACAGACAGGAGGGAATGATCTCTCC 1509
Db 131 TCCACGACATTCGGCTGAGCATATCTTTCTCTCTACAGACAGGAGGGAATTTGGCTCTCC 190

Qy 1510 ACATTGGTTCGAGACTTCTTTGTGATGAAAGAGGAGTCTTGCTCGTGAATTTTGACCTG 1569
Db 191 ACATTGGTTCGAGACTTCTTTGTGATGAAAGAGGAGTCTTGCTCGTGAATTTTGACTTG 250

Qy 1570 GGGTTCAGGAAATTCAGAAATGATAGTATGATGATGATGATGATGATGATGATGATG 1629
Db 251 GGGTTCAGGAAATTCAGAAATGATAGTATGATGATGATGATGATGATGATGATGATGATG 310

Qy 1630 GTCAAAATCACCCACATAGCAGGAGCATGATGATGATGATGATGATGATGATGATGATG 1689
Db 311 ATCAATATCACCCACATAGCAGGAGCATGATGATGATGATGATGATGATGATGATGATGATG 370

Qy 1690 CCATCAGTCTTTCGAACTCAACTTCTACAGCAATGGGCTGCTTTCATGTCATGCA 1748
Db 371 GCATCAGTCTTTCGAACTCAACTTCTACAGCAATGGGCTGCTTTCATGTCATGCA 429

RESULT 12
US-09-960-352-8821
; Sequence 8821, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 8821
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 38-LIB2809-028-Q1-E1-B6
US-09-960-352-8821

Query Match      13.6%; Score 337.6; DB 10; Length 407;
Best Local Similarity 90.9%; Pred. No. 9.4e-97;
Matches 371; Conservative 0; Mismatches 34; Indels 3; Gaps 1;

Qy 1330 GGTAGAGACAGTCCATTAAAGAGTCCAGAAATGCAACAGATGAATCCCTACGAGGAGG 1389
Db 3 GATACAGACAAGTCCATTAAAGAGTCCAGAAATGCAACAGATGAATCCCGCGCA---AGA 59

Qy 1390 TTGATTGCAATTCGGCTGAGCATATCTATTCACTGCTAGCAAGTCTTGCTGCCATTATG 1449
```

```
Db 60 CTGATTGCACATCTGGGCTGAGCACAATCTCTTCACTGCTAGCAAGTCTGCTGCTATTATG 119
Qy 1450 TCCACACACATTGTTGGCTTGGCTGCTCTCTACAGACAGGAGGGAATTTGATCTCTCC 1509
Db 120 TCCACGACATCTGTTGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 179
Qy 1510 ACATTGGTTCGAAAGACTTCTTTGTGATGAAAGAGGAAAGTCTTGCTCGTGAATTTTGACCTG 1569
Db 180 ACATTGGTTCGAAAGACTTCTTTGTGATGAAAGAGGAAAGTCTTGCTCGTGAATTTTGACTTG 239
Qy 1570 GGGTTCAGGAAATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1629
Db 240 GGGTTCAGGAAATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 299
Qy 1630 GTCAAAATCACCCACATAGCAGGAGCATGATGATGATGATGATGATGATGATGATGATGATG 1689
Db 300 ATCAATATCACCCACATAGCAGGAGCATGATGATGATGATGATGATGATGATGATGATGATG 359
Qy 1690 CCATCAGTCTTTCGAACTCAACTTCTACAGCAATGGGCTGCTTTCATGTC 1737
Db 360 CCATCAGTCTTTCGAACTCAACTTCTACAGCAATGGGCTGCTTTCATGTC 407

RESULT 13
US-09-833-381-1958/c
; Sequence 1958, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1958
; LENGTH: 868
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(868)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1958

Query Match      13.4%; Score 331.8; DB 10; Length 868;
Best Local Similarity 99.1%; Pred. No. 1.2e-94;
Matches 333; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2149 TTTATCACCTTCTTTACAGAGACTCCTTTGGGCTTTGCTGGAGGCTACAGCTCTGCTGCC 2208
Db 865 TTTATCACCTTCTTTACAGAGACTCCTTTGGGCTTTGCTGGAGGCTTANCGCTCTGCTGCC 806
Qy 2209 ATCTTTGTTTACACACTTTCAGTGGTCTGTTCCAGAACTGATGATCTGCAAAAGTTGCAC 2268
Db 805 ATCTTTGTTTACACACTTTCAGTGGTCTGTTCCAGAACTGATGATCTGCAAAAGTTGCAC 746
Qy 2269 AAATACCTTAATAACAGAAAGAAATGTTTCAGTATATGCTGAGAGTCCACATAT 2328
Db 745 AAATACCTTAATAACAGAAAGAAATGTTTCAGTATATGCTGAGAGTCCACATAT 686
Qy 2329 TGTCTTGTGAAGAAATGCTGTGAAATGTTTAAAGATATTTGGGTTTTCAGAGAGACCAA 2388
Db 685 TGTCTTGTGAAGAAATGCTGTGAAATGTTTAAAGATATTTGGGTTTTCAGAGAGACCAA 626
Qy 2389 CAAAAGAGAGTCTGTTTGTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGCAA 2448
Db 625 CAAAAGAGAGTCTGTTTGTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGCAA 566
Qy 2449 AAATCTCTAGAAATATTTCTGAGTCTGTTGCTGCTG 2484
```

Db 565 AAACCTTAGAATATATCTGAGTTTGTGGTGCTG 530

## RESULT 14

US-09-960-352-742/c  
; Sequence 742, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathalagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; TITLE OF INVENTION: NUCLEIC ACID AND FAT DEPOSITION  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960.352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 742  
; LENGTH: 418  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (395)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: 04-LIB2809-001-Q1-E1-A11  
US-09-960-352-742

Query Match 12.2%; Score 302; DB 10; Length 418;  
Best Local Similarity 84.1%; Pred. No. 2.4e-85;  
Matches 354; Conservative 0; Mismatches 61; Indels 6; Gaps 1;  
QY 1772 ATGAGTCTTGAACAGAGGGGACTGGGGGTCCACTAGCACCCACCTAACTGATCA 1831  
Db 418 ATGAGTCTTGAACAGAGGGGCGNCAGGAGACC-----CGGGTCTCCAGCTTGGTCA 365  
QY 1832 GCAGAGCAGCTGGTGGGAGGCGGCGGCTGTGCTACCTTCTTCCCAATGAAGGCA 1891  
Db 364 GCAGAGGAGCTGGTGGCAGAGCTGCCAGCTGTGCTATCTGCTCTCCAATGAAGGCA 305  
QY 1892 CCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAACAGTAGGAAAGTTTA 1951  
Db 304 CCATCTCTCTCCCTGCCAGACCTTTTACCATAATTTGCCATGAACAGTAGGCGGTTTA 245  
QY 1952 TCAGTAGGCATCTTTACAGTGGCAGAGCAGATGACAGGAGATATCAGTCTTAGTC 2011  
Db 244 TCAGTAGTTGCAATTTTATAGTTGCCGAGCAAGATAATCAGGAGATATCAGTCTTAGTT 185  
QY 2012 TTGCTGAGCAGCTGGGCAAGAAGCTTCCAGAACCTTTGTCTTGGAGAAGTGATGAAG 2071  
Db 184 CTTCCGAGCAGCTGGGCAAGAAGTTCCGGAACCTTTGTCTTGGAGAAGTGATGAAG 125  
QY 2072 AAGATGAAGCAGCTGACTTTGGGGGGAACAGGAGATTTCTACCTGAAGGTGAGCCAAT 2131  
Db 124 AAGATGAGACAGCGATTTTGGTGAGGAGCAACGAGATGCTACCTGAAGGTGAGCCAAT 65  
QY 2132 CCAAGGAGCACAGCAGTTTATACCTTCTTACAGAGATCTCTTGGGCTTTCTGAGG 2191  
Db 64 CCAAGGAGCACAGCAGTTTATACCTTCTTCTGAGAGGCTCTCTCGGCGCTCTGCTGAGG 5  
QY 2192 C 2192  
Db 4 C 4

## RESULT 15

US-09-814-353-16182/c  
; Sequence 16182, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John

; APPLICANT: Thompson, Pamela  
; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
; FILE REFERENCE: MRI-006B  
; CURRENT APPLICATION NUMBER: US/09/814.353  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/191,031  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: US 60/207,124  
; PRIOR FILING DATE: 2000-03-25  
; PRIOR APPLICATION NUMBER: US 60/211,940  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 60/216,820  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/220,661  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: US 60/257,672  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 22037  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16182  
; LENGTH: 379  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-814-353-16182  
Query Match 10.8%; Score 268.6; DB 13; Length 379;  
Best Local Similarity 96.6%; Pred. No. 1.1e-74;  
Matches 285; Conservative 0; Mismatches 9; Indels 1; Gaps 1;  
QY 1438 TGTGCCATTATGTCACACACATTTGGCTTCCTCTCTACAGACACAGGAGGGA 1497  
Db 379 TGTGCCATTATGTCACACACATTTGGCTTCCTCTCTACAGACACAGGAGGGA 320  
QY 1498 ATTGATCTCTCCACATTTGGTCGAAGACTTCTTTGTGATGAAAGAGAGAGTCTCTGGCTCT 1557  
Db 319 AATGATCTCTCCACATTTGGTCGAAGACTTCTTTGTGATGAAAGAGAGAGTCTCTGGCTCT 260  
QY 1558 GATTTTGACCTGGGTTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGCTG 1617  
Db 259 AATAATTGACCTGGGTTCTCAGGAAATTCAGAAACATGTAGTAATGCATGCCATACAGCTG 200  
QY 1618 CTGGGAAATTTGTTCACATCACCACACTAGCAGGAAACGATGAGTTTTTTTATCACCCCC 1677  
Db 199 CTGGGAAATTTGTTCACATCACCACACTAGCAGGAAACGATGAGTTTTTTTATCACCCCC 141  
QY 1678 AGCACAACCTGTCCCATCAGTCTTCGAACTCAATCTTCTACAGCAATGGGATCTTC 1732  
Db 140 AGCACAACCTGTCCCATCAGTCTTCGAACTCAATCTTCTACAGCAATGGGATCTTC 86

Search completed: January 10, 2004, 10:32:05  
Job time : 577.937 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 18:12:45 ; Search time 5946.75 Seconds  
(without alignments)  
17088.240 Million cell updates/sec

Title: US-09-935-290-3  
Perfect score: 2484  
Sequence: 1 atcgatgaatctgcactgac.....ttctgagtttttgtggtgctg 2484

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_mu:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_un:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rod:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hgto\_hum:\*  
40: em\_hgto\_mus:\*  
41: em\_hgto\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2484	100.0	2484	6	AX420455	AX420455 Sequence
2	2484	100.0	3003	6	AX420453	AX420453 Sequence
3	2482.4	99.9	2487	6	AX431391	AX431391 Sequence
4	2480.8	99.9	3210	6	AX451419	AX451419 Sequence
5	2479.2	99.8	3490	9	BC030783	BC030783 Homo sapi
6	2479.2	99.8	4918	9	HSN804404	HSN804404 Homo sapi
7	2477.6	99.7	6387	9	AL833093	AL833093 Homo sapi
8	2111.8	85.0	5679	9	HSN804372	HSN804372 Homo sapi
9	1984.4	79.9	5684	9	AB046780	AB046780 Homo sapi
10	1879.2	75.7	2646	6	AX401972	AX401972 Sequence
11	1879.2	75.7	2646	10	AF021348	AF021348 Rattus no
12	1860	74.9	3857	10	BC019201	BC019201 Mus muscu
13	1856.8	74.8	6634	10	MUS3GPAT	M77003 Mouse glyco
14	1803	72.6	2519	10	RNU36771	U36771 Rattus norv
15	267.4	10.8	196657	9	AL391986	AL391986 Human DNA
16	223.6	9.0	255430	2	AC099387	AC099387 Rattus no
17	218.2	8.8	197741	2	AC110206	AC110206 Mus muscu
18	173.8	7.0	206914	2	AC104911	AC104911 Mus muscu
19	142.6	5.7	4201	4	AF469047	AF469047 Bos tauru
20	131.6	5.3	147242	2	AC135152	AC135152 Rattus no
21	123.2	5.0	147242	2	AC135152	AC135152 Rattus no
22	109	4.4	2683	6	AX710670	AX710670 Sequence
23	109	4.4	2755	6	AX451420	AX451420 Sequence
24	107.8	4.3	2382	6	AX710668	AX710668 Sequence
25	107.8	4.3	2388	6	AX710666	AX710666 Sequence
26	107.8	4.3	2665	6	AX710664	AX710664 Sequence
27	103.8	4.2	3450	3	AF145635	AF145635 Drosophil
28	103.8	4.2	76299	2	AC019585	AC019585 Drosophil
29	103.8	4.2	172363	3	AC009388	AC009388 Drosophil
30	103.8	4.2	206190	3	AE003763	AE003763 Drosophil
31	95.4	3.8	539	11	G98364	G98364 S210P6250F
32	85.6	3.4	38145	3	C64847	U64847 Caenorhabdi
33	75	3.0	40532	3	CBRG36D18	AC084561 Caenorhab
34	64.8	2.6	301995	1	AE016779	AE016779 Pseudomon
35	61.4	2.5	10744	1	AE013559	AE013559 Yersinia
36	61.4	2.5	214050	1	AJ414142	AJ414142 Yersinia
37	59.2	2.4	300029	1	AE012554	AE012554 Xylella f
38	58	2.3	310029	1	AE016861	AE016861 Pseudomon
39	56.6	2.3	299338	1	AE016800	AE016800 Vibrio vu
40	56	2.3	11392	1	AE003940	AE003940 Xylella f
41	55	2.2	10802	1	AE015425	AE015425 Shigella
42	55	2.2	292497	1	AE016989	AE016989 Shigella
43	54.2	2.2	12847	1	AE012079	AE012079 Xanthomon
44	53.4	2.1	1141	6	AX083744	AX083744 Sequence
45	53	2.1	12739	1	AE004787	AE004787 Pseudomon

## ALIGNMENTS

RESULT 1				
AX420455				
LOCUS		2484 bp	DNA	linear
DEFINITION	Sequence 3 from Patent WO0216592.			
ACCESSION	AX420455			
VERSION	AX420455.1	GI:21524604		
KEYWORDS	.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1			
AUTHORS	Kapeller-Libermann,R. and Logan,T.J..			
TITLE	Atcr-1, a human acyltransferase and uses thereof			
JOURNAL	Patent: WO 0216592-A 3 28-FEB-2002:			



Db 1981 CACGATGACCAAGAGATATCAGTCTCTAGTCTTGCTGAGCAGCAGTGGCAAGAAGCTT 2040  
Qy 2041 CCAGAACCTTTGCTTTGGAGAAGTGATGAAGAAGATGAAGACAGTGAATTTGGGGAGGAA 2100  
Db 2041 CCAGAACCTTTGCTTTGGAGAAGTGATGAAGAAGATGAAGACAGTGAATTTGGGGAGGAA 2100  
Qy 2101 CAGCGAGATGCTACCTGAGGTGAGCAATCCAAAGGAGCACCAGCAGTGTATCACCTTC 2160  
Db 2101 CAGCGAGATGCTACCTGAGGTGAGCAATCCAAAGGAGCACCAGCAGTGTATCACCTTC 2160  
Qy 2161 TTACAGAGACTCTCTGGGCTTTGCTGAGGCTACAGCTGCTGCGCATCTTTGTTTAC 2220  
Db 2161 TTACAGAGACTCTCTGGGCTTTGCTGAGGCTACAGCTGCTGCGCATCTTTGTTTAC 2220  
Qy 2221 AACTTCAGTGTCTGTTCCAGAACTCTGAGTATCTGCAAAAGTTCACAAATACCTAATA 2280  
Db 2221 AACTTCAGTGTCTGTTCCAGAACTCTGAGTATCTGCAAAAGTTCACAAATACCTAATA 2280  
Qy 2281 ACCAGAACAGAAAGAAATGTTGAGATATGCTGAGAGTGCACATATGCTTGTGAAG 2340  
Db 2281 ACCAGAACAGAAAGAAATGTTGAGATATGCTGAGAGTGCACATATGCTTGTGAAG 2340  
Qy 2341 AATGCTGTGAAATCTTTAAGATATGCTGAGAGTGCACATATGCTTGTGAAG 2400  
Db 2341 AATGCTGTGAAATCTTTAAGATATGCTGAGAGTGCACATATGCTTGTGAAG 2400  
Qy 2401 TCTGTTTGTAGAACTGAGCAGCACTTTCTACCTCAATCAACCGACAAACTTCTAGAA 2460  
Db 2401 TCTGTTTGTAGAACTGAGCAGCACTTTCTACCTCAATCAACCGACAAACTTCTAGAA 2460  
Qy 2461 TATATCTGAGTTTGTGCTGCTG 2484  
Db 2461 TATATCTGAGTTTGTGCTGCTG 2484

RESULT 2  
AX420453  
LOCUS AX420453 3003 bp DNA linear PAT 18-JUN-2002  
DEFINITION Sequence 1 from Patent WO0216592.  
ACCESSION AX420453  
VERSION AX420453.1 GI:21524602  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
Kapeller-Libermann, R. and Logan, T.J.  
Atr-1, a human acyltransferase and uses thereof  
Patent: WO 0216592-A 1 28-FEB-2002;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
source  
Location/Qualifiers  
1..3003  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
341..2827  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD35486.1"  
/db\_xref="GI:21524603"  
/translation="WDSEALTLGIDVSLPHSSEYSVGRCKHTSEWBGCPPTVP  
RSATLKWESLSMRKPPVGCYSCYSPQWDKFPNPSISLGLRNVYINETHRR  
GWLARRLSYLFIERDVKHGFNFAINTVNLNRSVQEAIAEVAELNPDGSAQQS  
KANVKKAKRILOEAVATVPAMIRLTGWLKLFNSFFWFOIHKQLEWKAAT  
ETNLPFLPHVRSHIDYLLFLFCHNIKAPYIAGSNLNIPIFSLIHLGGLFI  
RRLEDPDGRKDVYLLHLLHVELLRQOFLIEIIEGTRSRSGTSCARAGLVS  
VVDLTSTNVI PDILLI PVGISYDRI IEHYNGEOLGPKKNESLWSVARGVIRMLRN  
YGVGVDPQAPFSLKEVLESQSPVSNALLSLEALLPAILPSRPSRADEGRDTSN  
ESRNATESRRRIANLAELHILFASKSASTMTHIVACLLLYRHQGLDLSLVED  
FFVWKBEVLARDFDLPGSGNSEDVVMHAIQLLGNCTVITHTSRNDEFITPSTVPSV  
FELNFSYNGVLHVIMEAIIACSLYAVLNKRGLGPTSTPPNLSIQBQLVRKAASLCY

LLSNEGTSILPQCTFYQVCHETVGKFIQYGLTVAEHDDQEDISPSLAEOQWKKLPE  
NFSRSEDESDSFGSEQRDCLVKVQSKEHQOIFITFLRLLGLPLEAYSSAIFVH  
NFGPVPPEYVLOKLHKYLTIRTRNVAVAESATYCLVKNVAVKMFQDVGFKETKOK  
RVSVLELSSFTFLPCNQKLELXILTSFVL" 2 others  
BASE COUNT 826 a 683 c 698 g 794 t  
ORIGIN  
Query Match 100.0%; Score 2484; DB 6; Length 3003;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGATGAATCTGACATGACCCCTTGGTACAAATAGATGTTTCTATCTGCCACATTCATCA 60  
Db 341 ATGATGAATCTGACATGACCCCTTGGTACAAATAGATGTTTCTATCTGCCACATTCATCA 400  
Qy 61 GAATACAGTGTGTCGATGTAAGCACACAGTGCAGAAATGGGCTGAGTGGCTTTAGA 120  
Db 401 GAATACAGTGTGTCGATGTAAGCACACAGTGCAGAAATGGGCTGAGTGGCTTTAGA 460  
Qy 121 CCCACCGTCTTCAGATCTGCAACTTTTAAATGGAAGAAAGCCCTAATGAGTCGGAAGG 180  
Db 461 CCCACCGTCTTCAGATCTGCAACTTTTAAATGGAAGAAAGCCCTAATGAGTCGGAAGG 520  
Qy 181 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCAGAGCTGGGCAAAATTTTCAAC 240  
Db 521 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCAGAGCTGGGCAAAATTTTCAAC 580  
Qy 241 CCCAGTATCCGCTTTTGGGTTTGGGAATGTTATTTATATCAATGAATCAACACAGA 300  
Db 581 CCCAGTATCCGCTTTTGGGTTTGGGAATGTTATTTATATCAATGAATCAACACAGA 640  
Qy 301 CACCGCGGATGGCTTGCAAGACGCTTTCTTACCTCTTTTATTTTCAAGAGCAGAGTGTG 360  
Db 641 CACCGCGGATGGCTTGCAAGACGCTTTCTTACCTCTTTTATTTTCAAGAGCAGAGTGTG 700  
Qy 361 CATAAGGCGATGTTTGGCAACCAATGTGACTGAAATGTCTGAAACAGCAGTGTAGTACAA 420  
Db 701 CATAAGGCGATGTTTGGCAACCAATGTGACTGAAATGTCTGAAACAGCAGTGTAGTACAA 760  
Qy 421 GAGCAATTCAGAAAGTGGCTGCTGAATTAACCCCTGATGTTCTGCCAGCAGCAATCA 480  
Db 761 GAGCAATTCAGAAAGTGGCTGCTGAATTAACCCCTGATGTTCTGCCAGCAGCAATCA 820  
Qy 481 AAAGCCGTAAACAAAGTGAAGAAAGAAAGAAAGTCTTCAAGAAATGTTTGGCCACT 540  
Db 821 AAAGCCGTAAACAAAGTGAAGAAAGAAAGAAAGTCTTCAAGAAATGTTTGGCCACT 880  
Qy 541 GTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db 881 GTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940  
Qy 601 TTTTGGAAACATTCAAATTCAGAAAGTCAACTGAGATGTTTAAAGCTGCAATGTAGACG 660  
Db 941 TTTTGGAAACATTCAAATTCAGAAAGTCAACTGAGATGTTTAAAGCTGCAATGTAGACG 1000  
Qy 661 AATTTGCGCTCTGTTTCTTACAGTTCATAGATCCCATATATGACTATCTGCTGCTGCTGCT 720  
Db 1001 AATTTGCGCTCTGTTTCTTACAGTTCATAGATCCCATATATGACTATCTGCTGCTGCTGCT 1060  
Qy 721 TTCAATCTCTTCTGCCATAACATCAAGACCAATGCTTCAAGCAATTAATCTCAAC 780  
Db 1061 TTCAATCTCTTCTGCCATAACATCAAGACCAATGCTTCAAGCAATTAATCTCAAC 1120  
Qy 781 ATCCCAATCTTCACTACCTGATCCATAAGCTTGGGGCTTCTTTCATACGACGAAGGCTC 840  
Db 1121 ATCCCAATCTTCACTACCTGATCCATAAGCTTGGGGCTTCTTTCATACGACGAAGGCTC 1180  
Qy 841 GATGAAACACAGATGACGGAAGATGTTCTCTATAGAGCTTGTCTTCAATGCGGCATATA 900  
Db 1181 GATGAAACACAGATGACGGAAGATGTTCTCTATAGAGCTTGTCTTCAATGCGGCATATA 1240  
Qy 901 GTTGAATTTACTTCGACAGCAGCAATTTCTTGAGATCTTCTGGAAGGACACAGCTTCTAGG 960  
Db 961 GTTGAATTTACTTCGACAGCAGCAATTTCTTGAGATCTTCTGGAAGGACACAGCTTCTAGG

```
Db 1241 GTTGAATTACTTCGACAGCAGCAATCTTGGAGATCTTCTGGAAGGCACACGCTTCAGG 1300
Qy 961 AGTGGAAAAAACCCTCTTGTGCTCGGGCAGGACTTTTGTGTCAGTGTGGTAGATACCTCTGTCT 1020
Db 1301 AGTGGAAAAAACCCTCTTGTGCTCGGGCAGGACTTTTGTGTCAGTGTGGTAGATACCTGTCT 1360
Qy 1021 ACCAATGTCATCCAGACATCTTTGATAATACCTGTTGGAATCTCTATGATCGCATTTATC 1080
Db 1361 ACCAATGTCATCCAGACATCTTTGATAATACCTGTTGGAATCTCTATGATCGCATTTATC 1420
Qy 1081 GARAGTCACATCAATGCTGAACAACCTGGGCAACCTTAAGAAGATGAGAGCCCTGTGAGT 1140
Db 1421 GAAGGTCACTACAAATGGTGAACAACCTGGGCAACCTTAAGAAGATGAGAGCCCTGTGAGT 1480
Qy 1141 GTAGCAAGAGGTGTTATTGAATGTTTACGAAAAAACTATGTTGTGTGTCGAGTGGATTTT 1200
Db 1481 GTAGCAAGAGGTGTTATTGAATGTTTACGAAAAAACTATGTTGTGTGTCGAGTGGATTTT 1540
Qy 1201 GCACAGCCATTTTCTTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCT 1260
Db 1541 GCACAGCCATTTTCTTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCT 1600
Qy 1261 CTACTTTCCCTGGAGCAGAGGTTGTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320
Db 1601 CTACTTTCCCTGGAGCAGAGGTTGTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1660
Qy 1321 GCTGATGAAGGTAGAGACAGTCCTAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380
Db 1661 GCTGATGAAGGTAGAGACAGTCCTAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1720
Qy 1381 CGAAGGAGGTGATTTGCAAAATCTGGCTGAGCATATTTCTATCTGCTAGCAAGTCTGT 1440
Db 1721 CGAAGGAGGTGATTTGCAAAATCTGGCTGAGCATATTTCTATCTGCTAGCAAGTCTGT 1780
Qy 1441 GCATATATGTCACACACATTTGCTGCTGCTGCTCTCTCTACAGACACAGGCGAGGAAT 1500
Db 1781 GCATATATGTCACACACATTTGCTGCTGCTGCTCTCTCTCTACAGACACAGGCGAGGAAT 1840
Qy 1501 GATCTCTCCACATTTGCTGGAAGACTCTTTGTGATGAAGAGGAAGTCTCGGCTCGTGAT 1560
Db 1841 GATCTCTCCACATTTGCTGGAAGACTCTTTGTGATGAAGAGGAAGTCTCGGCTCGTGAT 1900
Qy 1561 TTTGACTGGGGTCTCAGGAAATTCAGAAATGTAGTAATGCAATGCATACAGCTGCTG 1620
Db 1901 TTTGACTGGGGTCTCAGGAAATTCAGAAATGTAGTAATGCAATGCATACAGCTGCTG 1960
Qy 1621 GGAATTTGTCACATCACCACACTAGCAGGAAACGATGAGTTTTTTATCACCACCCAGC 1680
Db 1961 GGAATTTGTCACATCACCACACTAGCAGGAAACGATGAGTTTTTTATCACCACCCAGC 2020
Qy 1681 ACAACTGTCCCATCAGTCTTCGAACTCAACTTCTACAGCAATGGGTAATGCTGCTTT 1740
Db 2021 ACAACTGTCCCATCAGTCTTCGAACTCAACTTCTACAGCAATGGGTAATGCTGCTTT 2080
Qy 1741 ATCATGGAGGCCATCATAGCTTTCAGAGCTTTATGCAAGTCTGTAACAGAGGGAGCTGGG 1800
Db 2081 ATCATGGAGGCCATCATAGCTTTCAGAGCTTTATGCAAGTCTGTAACAGAGGGAGCTGGG 2140
Qy 1801 GGTCCCACTAGCACCCACCTAATCCTGATCAGCAGGAGCAGCTGGTGGGAAGGGCGCC 1860
Db 2141 GGTCCCACTAGCACCCACCTAATCCTGATCAGCAGGAGCAGCTGGTGGGAAGGGCGCC 2200
Qy 1861 AGCTGTGCTACCTCTCTCCATGAGGACCATCTCAGTCTGCTGCTGCGCAGATTTTAC 1920
Db 2201 AGCTGTGCTACCTCTCTCCATGAGGACCATCTCAGTCTGCTGCTGCGCAGATTTTAC 2260
Qy 1921 CAAAGTCTGCCATGAACAGTAGGAAGTTTATCCAGTATGGCATTTCTACAGTGGCAGAG 1980
Db 2261 CAAAGTCTGCCATGAACAGTAGGAAGTTTATCCAGTATGGCATTTCTACAGTGGCAGAG 2320
Qy 1981 CACGATGACGAGGAAGATATCAGTCTCTGCTGCTGAGCAGCAGTGGGCAAGAGAGCTT 2040
Db 2321 CACGATGACGAGGAAGATATCAGTCTCTGCTGCTGAGCAGCAGTGGGCAAGAGAGCTT 2380
```

```
Qy 2041 CCAGAACCTTTGTCTTGGAGAAAGTGATGAAGAAGATGAAGACAGTCTTCTGGGGAGGAA 2100
Db 2381 CCAGAACCTTTGTCTTGGAGAAAGTGATGAAGAAGATGAAGACAGTCTTCTGGGGAGGAA 2440
Qy 2101 CAGCGAGATTGCTACTCTGAAGGTGAGCCAAATCCAGGAGCACCAGCAGTTTATCAGCTTC 2160
Db 2441 CAGCGAGATTGCTACTCTGAAGGTGAGCCAAATCCAGGAGCACCAGCAGTTTATCAGCTTC 2500
Qy 2161 TTACAGAGACTCTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2220
Db 2501 TTACAGAGACTCTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2560
Qy 2221 AACTTCAGTGGTCTCTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAATA 2280
Db 2561 AACTTCAGTGGTCTCTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAATA 2620
Qy 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATTTGTTGTGAAG 2340
Db 2621 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATTTGTTGTGAAG 2680
Qy 2341 AATGCTGTGAAAATGTTTAAAGGATATTTGGGGTTTCAAGGAGACCAAAACAAAGAGAGTG 2400
Db 2681 AATGCTGTGAAAATGTTTAAAGGATATTTGGGGTTTCAAGGAGACCAAAACAAAGAGAGTG 2740
Qy 2401 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAAACTTCTAGAA 2460
Db 2741 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAAACTTCTAGAA 2800
Qy 2461 TATATTCTCAGTTTGTGTGCTG 2484
Db 2801 TATATTCTCAGTTTGTGTGCTG 2824
```

## RESULT 3

```
AX431391 LOCUS AX431391 2487 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 24 from Patent WO0240666.
ACCESSION AX431391
VERSION AX431391.1 GI:21656238
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Enkins,D.K., Winther,M.D., Haardt,M., Goldberg,Y.P., Nwaka,S.O.,
Ponton,A., Allen,S.J., de Antueno,R.J. and Knickle,L.C.
TITLE Fat regulated genes, uses thereof, and compounds for modulating
same
JOURNAL Patent: WO 0240666-A 24 23-MAY-2002;
XENON GENETICS INC (CA)
FEATURES
source
Location/Qualifiers
1..2487
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 706 a 548 c 570 g 663 t
ORIGIN
```

```
Query Match 99.9%; Score 2482.4; DB 6; Length 2487;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2483; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 ATGGATGAATCTGCACTGACCTTGGTCAATAGATGTTTCTTATCTGCCAATTCATCA 60
Db 1 ATGGATGAATCTGCACTGACCTTGGTCAATAGATGTTTCTTATCTGCCAATTCATCA 60
Qy 61 GAATACAGTGTGGTTCGATGAAGCACACAGTGGGAATGGGTGAGTGGCTTTTGA 120
Db 61 GAATACAGTGTGGTTCGATGAAGCACACAGTGGGAATGGGTGAGTGGCTTTTGA 120
Qy 121 CCCACGCTTTCAGATCTCAACTTTTAAATGGAAGAAAGCCTAATAGTTCGGAAGG 180
```

Db 121 CCAACATCTTCAGATCTGCAACTTTAAATGAAAGAAAGCCTAATGAGTCGGAAAGG 180  
Qy 181 CCAATTTGTTGGAAGATGTTGTTACTCTCGCACTCCCCAGAGCTGGGACAAAATTTTCAAC 240  
Db 181 CCAATTTGTTGGAAGATGTTGTTACTCTCGCACTCCCCAGAGCTGGGACAAAATTTTCAAC 240  
Qy 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAACTCACACAAGA 300  
Db 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAACTCACACAAGA 300  
Qy 301 CACCCGGATGCTTTGCAAGAGCGCTTCTTACAGTTCTTTTATTTCAAGAGCGAGATGTG 360  
Db 301 CACCCGGATGCTTTGCAAGAGCGCTTCTTACAGTTCTTTTATTTCAAGAGCGAGATGTG 360  
Qy 361 CATAAGGGCATGTTTGCCACCAATGTGACTGAAAATGTGCTGAAACAGCAGTAGAGTACAA 420  
Db 361 CATAAGGGCATGTTTGCCACCAATGTGACTGAAAATGTGCTGAAACAGCAGTAGAGTACAA 420  
Qy 421 GAGGCAATTCGAAAGTGGCTGCTGAAATTAACCCCTGATGTTCTGCCAGCAGCAATCA 480  
Db 421 GAGGCAATTCGAAAGTGGCTGCTGAAATTAACCCCTGATGTTCTGCCAGCAGCAATCA 480  
Qy 481 AAAGCCGTTTAAACAAGTGAAGAAAGAAAGCTAAAGGATTTCTCAAGAAATGGTTGCCACT 540  
Db 481 AAAGCCGTTTAAACAAGTGAAGAAAGAAAGCTAAAGGATTTCTCAAGAAATGGTTGCCACT 540  
Qy 541 GTCTCACCGGCAATGATCAGACTGAGCTGGGTGGGTGCTGCTAAACTGTTCAACAGCTTC 600  
Db 541 GTCTCACCGGCAATGATCAGACTGAGCTGGGTGGGTGCTGCTAAACTGTTCAACAGCTTC 600  
Qy 601 TTTTGGAACTTCAAATTCACAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACG 660  
Db 601 TTTTGGAACTTCAAATTCACAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACG 660  
Qy 661 AATTTGCCGCTCTGTTTCTACCAAGTTCATGATCCCAATATGACTATCTGCTGCTCACT 720  
Db 661 AATTTGCCGCTCTGTTTCTACCAAGTTCATGATCCCAATATGACTATCTGCTGCTCACT 720  
Qy 721 TTCAATCTCTCTGCGCATAACATCAAGCACCATACTGCTTCAAGGCAATAATCTCAAC 780  
Db 721 TTCAATCTCTCTGCGCATAACATCAAGCACCATACTGCTTCAAGGCAATAATCTCAAC 780  
Qy 781 ATCCCAATCTTCAGTACCTTGATCCAATGCTTTGGGGCTTCTTCATACGACGAGGCTC 840  
Db 781 ATCCCAATCTTCAGTACCTTGATCCAATGCTTTGGGGCTTCTTCATACGACGAGGCTC 840  
Qy 841 GATGAAACACAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db 841 GATGAAACACAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Qy 901 GTTGAATTTACTTCGACAGCAGCAATTCCTTGAGATCTTCTTGGAAGGCACACGTTCTAGG 960  
Db 901 GTTGAATTTACTTCGACAGCAGCAATTCCTTGAGATCTTCTTGGAAGGCACACGTTCTAGG 960  
Qy 961 AGTGGAAAAACCTCTTGCTCGGCGAGGACTTTTGTCAAGTTGTTGGTAGACTCTGCT 1020  
Db 961 AGTGGAAAAACCTCTTGCTCGGCGAGGACTTTTGTCAAGTTGTTGGTAGACTCTGCT 1020  
Qy 1021 ACCAATGTCATCCAGACATCTTGATATACTGTTGGAATCTCTATGATGCAATATC 1080  
Db 1021 ACCAATGTCATCCAGACATCTTGATATACTGTTGGAATCTCTATGATGCAATATC 1080  
Qy 1081 GAAGTCTACTACAATGGTGAACTGGGCAAACTTAAGAGAAATGAGAGCTGTGGAGT 1140  
Db 1081 GAAGTCTACTACAATGGTGAACTGGGCAAACTTAAGAGAAATGAGAGCTGTGGAGT 1140  
Qy 1141 GTAGCAAGAGGTTTATTAGATGTTACGAAAAAACTATGGTTGTTGTCGAGTGGATTTT 1200  
Db 1141 GTAGCAAGAGGTTTATTAGATGTTACGAAAAAACTATGGTTGTTGTCGAGTGGATTTT 1200  
Qy 1201 GCACAGCCATTTCTTAAAGGAATTTTAAAGGCCAAAGTCAGAAACCGGTGTCTGCT 1260

Db 1201 GCACAGCCATTTTCTTAAAGGAATATTTTAAAGGCCAAAGTCAGAAACCGGTGTCTGCT 1260  
Qy 1261 CTACTTTCCCTGGAGCAAGCGTTGTTACCACTATATCTTCTTCAAGACCCAGTGTGCT 1320  
Db 1261 CTACTTTCCCTGGAGCAAGCGTTGTTACCACTATATCTTCTTCAAGACCCAGTGTGCT 1320  
Qy 1321 GCTGATGAGGTAGAGACAGTCCATTAATAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db 1321 GCTGATGAGGTAGAGACAGTCCATTAATAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Qy 1381 CGAAGGAGGTGATTTGCAAAATCTGGCTGAGCATATCTATTTCACTGTAGCAAGTCTGT 1440  
Db 1381 CGAAGGAGGTGATTTGCAAAATCTGGCTGAGCATATCTATTTCACTGTAGCAAGTCTGT 1440  
Qy 1441 GCCATTAATGTCACACACATTTGCTGCTGCTCTCTACAGACACAGGACGAGGAAT 1500  
Db 1441 GCCATTAATGTCACACACATTTGCTGCTGCTCTCTCTACAGACACAGGACGAGGAAT 1500  
Qy 1501 GATCTCTCCATTTGGTTCGAAGACTTCTTTGTGATGAAAGGAAGTCTTGCTGCTGAT 1560  
Db 1501 GATCTCTCCATTTGGTTCGAAGACTTCTTTGTGATGAAAGGAAGTCTTGCTGCTGAT 1560  
Qy 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAATTCAGAAATTCAGTAAATGTCATACAGCTGCTG 1620  
Db 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAATTCAGAAATTCAGTAAATGTCATACAGCTGCTG 1620  
Qy 1621 GGAATTTGTGTCACAACTCACCACACTAGCAGGAACGATGAGTTTTTTTATCACCCCGAGC 1680  
Db 1621 GGAATTTGTGTCACAACTCACCACACTAGCAGGAACGATGAGTTTTTTTATCACCCCGAGC 1680  
Qy 1681 ACAACTGCTCCATCAGTCTTCGAACTCACTTCTACAGCAATGGGGTACTTCATGTCTTT 1740  
Db 1681 ACAACTGCTCCATCAGTCTTCGAACTCACTTCTACAGCAATGGGGTACTTCATGTCTTT 1740  
Qy 1741 ATCATGAGGCCATCATAGCTTGCAGCTTTATGCAAGTCTGTAACAAGAGGGGACTGGGG 1800  
Db 1741 ATCATGAGGCCATCATAGCTTGCAGCTTTATGCAAGTCTGTAACAAGAGGGGACTGGGG 1800  
Qy 1801 GGTCCTACTAGCAGCCACCTAACCTGATCAGCAGGAGCAGCTGGTCCGGAAGCGGCC 1860  
Db 1801 GGTCCTACTAGCAGCCACCTAACCTGATCAGCAGGAGCAGCTGGTCCGGAAGCGGCC 1860  
Qy 1861 AGCTGTGCTACCTTCTCTCCAATGAAGGACCACTCTCACTGCTTCCGACAGACATTTTAC 1920  
Db 1861 AGCTGTGCTACCTTCTCTCCAATGAAGGACCACTCTCACTGCTTCCGACAGACATTTTAC 1920  
Qy 1921 CAAGTCTGCCATGAAACAGTAGGAAAGTATCAAGTATGGCATTTCTTACGTGGCAGAG 1980  
Db 1921 CAAGTCTGCCATGAAACAGTAGGAAAGTATCAAGTATGGCATTTCTTACAGTGGCAGAG 1980  
Qy 1981 CACGATCAGCAGGAAGATATCAGTCTAGTCTGCTGAGCAGCAGTGGGCAAGAGCTT 2040  
Db 1981 CACGATCAGCAGGAAGATATCAGTCTAGTCTGCTGAGCAGCAGTGGGCAAGAGCTT 2040  
Qy 2041 CCAGAACTTTGCTTCCGAGAAAGTATGATGAAGAAAGTCAAGACAGTGTCTTTGGGAGGAA 2100  
Db 2041 CCAGAACTTTGCTTCCGAGAAAGTATGATGAAGAAAGTCAAGACAGTGTCTTTGGGAGGAA 2100  
Qy 2101 CAGCAGATTCCTACTGAAAGGTGAGCCAAATCCAAGGAGCAGCAGAGTGTATCACCTTC 2160  
Db 2101 CAGCAGATTCCTACTGAAAGGTGAGCCAAATCCAAGGAGCAGCAGAGTGTATCACCTTC 2160  
Qy 2161 TTACAGAGACTCTTGGGCTTGGTGGAGCCCTACAGCTCTGCTGCAATCTTTGTTTAC 2220  
Db 2161 TTACAGAGACTCTTGGGCTTGGTGGAGCCCTACAGCTCTGCTGCAATCTTTGTTTAC 2220  
Qy 2221 AACTTCAAGTGTCTGTTTCCAGAACTGAGTATCTGCAAAAGTTCGCAAAATCCTAATA 2280  
Db 2221 AACTTCAAGTGTCTGTTTCCAGAACTGAGTATCTGCAAAAGTTCGCAAAATCCTAATA 2280  
Qy 2281 ACCAGAAACAGAAAGAAATGTTGCAAGTATGCTGAGAGTGCCACATATGTTCTTGTGAAG 2340  
Db 2281 ACCAGAAACAGAAAGAAATGTTGCAAGTATGCTGAGAGTGCCACATATGTTCTTGTGAAG 2340



QY 2341 AATGCTGTGAAATGTTTAAAGATATTGGGTTTCAAGGAGACCACAAACAAAGAGAGTG 2400  
DB 2341 AATGCTGTGAAATGTTTAAAGATATTGGGTTTCAAGGAGACCACAAACAAAGAGAGTG 2400  
QY 2401 TCTGTTTAAAGTGTAGAGCAGCCTTTCTACCTCAATGCAACGACAAACAACTTTAGAA 2460  
DB 2401 TCTGTTTAAAGTGTAGAGCAGCCTTTCTACCTCAATGCAACGACAAACAACTTTAGAA 2460  
QY 2461 TATATCTGAGTTTGTGGTGCTG 2484  
DB 2461 TATATCTGAGTTTGTGGTGCTG 2484

RESULT 4  
AX451419  
LOCUS AX451419 3210 bp DNA linear PAT 03-JUL-2002  
DEFINITION Sequence 26 from Patent WO0226950.  
ACCESSION AX451419  
VERSION AX451419.1 GI:21698428  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Lal, P.G., Tang, Y.T., Yue, H., Burford, N., Gandhi, A.R., Warren, B.A.,  
Yao, M.G., Tribouley, C.M., Baughn, M.R., Lee, E.A., Hafalia, A.J.,  
Lu, Y., Griffin, J.A., Sanjanwala, M.S. and Ding, L.  
Transferrases  
TITLE Patent: WO 0226950-A 26 04-APR-2002;  
JOURNAL Incyte Genomics, Inc. (US)  
FEATURES  
source location/Qualifiers  
1..3210  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/notes="Incyte ID No: 1517230CB1"  
BASE COUNT 896 a 721 c 735 g 858 t  
ORIGIN

Query Match 99.9%; Score 2480.8; DB 6; Length 3210;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2482; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGATGAATCTGCATGACCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 60  
DB 338 ATGGATGAATCTGCATGACCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 397  
QY 61 GAATACAGTCTTGGTGTAGTAAAGCACAAAGTGGGATGGGTTGAGTGTGGCTTTAGA 120  
DB 398 GAATACAGTGTGGTGTAGTAAAGCACAAAGTGGGATGGGTTGAGTGTGGCTTTAGA 457  
QY 121 CCCACCGCTTTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAATAGTTCGGAAGG 180  
DB 458 CCCACCATCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAATAGTTCGGAAGG 517  
QY 181 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCAGAGCTGGGACAAATTTTCAAC 240  
DB 518 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCAGAGCTGGGACAAATTTTCAAC 577  
QY 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTTATATCAATGAACCTCACACAAGA 300  
DB 578 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTTATATCAATGAACCTCACACAAGA 637  
QY 301 CACCGGGATGGCTTGCAGACGCGCTTTCTTACGTTCTTTTATTCAGAGCGAGATGTG 360  
DB 638 CACCGGGATGGCTTGCAGACGCGCTTTCTTACGTTCTTTTATTCAGAGCGAGATGTG 697  
QY 361 CATAAGGCGATGTTTGGCCCAATGTGACTGGAATAATGTGCTGAACAGCAGTAGAGTACAA 420  
DB 698 CATAAGGCGATGTTTGGCCCAATGTGACTGGAATAATGTGCTGAACAGCAGTAGAGTACAA 757

QY 421 GAGGCAATTGCAGAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCAGAGCAATCA 480  
DB 758 GAGGCAATTGCAGAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCAGAGCAATCA 817  
QY 481 AAAGCGTTTAAACAAAGTGAAGAAAGCTAAAGGATTTCTTCAGAAATGTTGCCACT 540  
DB 818 AAAGCGTTTAAACAAAGTGAAGAAAGCTAAAGGATTTCTTCAGAAATGTTGCCACT 877  
QY 541 GTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTAAACCTGTTCAACAGCTTC 600  
DB 878 GTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTAAACCTGTTCAACAGCTTC 937  
QY 601 TTTTGGAAACATTTCAAATTCACAAAGGTCAAATTGAGATGGTTAAAGCTGCAACTGAGACG 660  
DB 938 TTTTGGAAACATTTCAAATTCACAAAGGTCAAATTGAGATGGTTAAAGCTGCAACTGAGACG 997  
QY 661 AATTTGCGCGTCTGTTTCTACAGTTCATAGATCCCATATTGACTATCTGCTGCTCACT 720  
DB 998 AATTTGCGCGTCTGTTTCTACAGTTCATAGATCCCATATTGACTATCTGCTGCTCACT 1057  
QY 721 TTCAATTTCTTCTGCGCATAAACATCAAGCACCATACATTGCTTCAGGCAATATCTCAAC 780  
DB 1058 TTCAATTTCTTCTGCGCATAAACATCAAGCACCATACATTGCTTCAGGCAATATCTCAAC 1117  
QY 781 ATCCCAATCTTCAGTACCTTGTATCCATAGCTTGGGGCTTCTTCATACGACGAAGGCTC 840  
DB 1118 ATCCCAATCTTCAGTACCTTGTATCCATAGCTTGGGGCTTCTTCATACGACGAAGGCTC 1177  
QY 841 GATGAACACAGATGGAGCGAAAGATGTTCTATAGAGCTTGTCTCATGGGCATATA 900  
DB 1178 GATGAACACAGATGGAGCGAAAGATGTTCTATAGAGCTTGTCTCATGGGCATATA 1237  
QY 901 GTTGAATTTACTTCGACAGCAGCAATTTCTTGGAGATCTTCTCGAAGGACACGTTCTAGG 960  
DB 1238 GTTGAATTTACTTCGACAGCAGCAATTTCTTGGAGATCTTCTCGAAGGACACGTTCTAGG 1297  
QY 961 AGTGGAAAAACCTCTTGTGCTCGGCGAGGACTTTTGTTCAGTTGTGGTAGATACTCTGTCT 1020  
DB 1298 AGTGGAAAAACCTCTTGTGCTCGGCGAGGACTTTTGTTCAGTTGTGGTAGATACTCTGTCT 1357  
QY 1021 ACCAATGTCTATCCAGACATCTTGTATATACCTGTTGGAATCTCTATGATCGCATATC 1080  
DB 1358 ACCAATGTCTATCCAGACATCTTGTATATACCTGTTGGAATCTCTATGATCGCATATC 1417  
QY 1081 GAAGGTCACTCAATAGTGAACAACTGGGCAAACTTAAGAAAGTGAAGCTGTGGAGT 1140  
DB 1418 GAAGGTCACTCAATAGTGAACAACTGGGCAAACTTAAGAAAGTGAAGCTGTGGAGT 1477  
QY 1141 GTAGCAAGAGGTGTTATTAGAAATGTTTACGAAAAAATCTATGTTGTCTCCGAGTGGATTT 1200  
DB 1478 GTAGCAAGAGGTGTTATTAGAAATGTTTACGAAAAAATCTATGTTGTCTCCGAGTGGATTT 1537  
QY 1201 GCACAGCCATTTTCTTAAAGGAATATTATTAAGCAAAAGTCAAGAACCGGTGTCTGCT 1260  
DB 1538 GCACAGCCATTTTCTTAAAGGAATATTATTAAGCAAAAGTCAAGAACCGGTGTCTGCT 1597  
QY 1261 CTACTTTCCCTCGAGCAAGCGTTGTACAGCTATCTTCTTCAAGCCAGTGTATGCT 1320  
DB 1598 CTACTTTCCCTCGAGCAAGCGTTGTACAGCTATCTTCTTCAAGCCAGTGTATGCT 1657  
QY 1321 GCTGATGAAGGTAGAGACACGTCCTAATATAGTCCAGAAATGCAACAGATGAATCCCTTA 1380  
DB 1658 GCTGATGAAGGTAGAGACACGTCCTAATATAGTCCAGAAATGCAACAGATGAATCCCTTA 1717  
QY 1381 CGAAGAGGTTGATTTGCAAAATCTGGCTGAGCATATTTCTATTCTGCTAGCAAGTCTGCT 1440  
DB 1718 CGAAGAGGTTGATTTGCAAAATCTGGCTGAGCATATTTCTATTCTGCTAGCAAGTCTGCT 1777  
QY 1441 GGCATTTATGTCACACATTTGCTGCTGCTGCTCTCTACAGACACAGGCGAGGAAT 1500  
DB 1778 GGCATTTATGTCACACATTTGCTGCTGCTGCTCTCTACAGACACAGGCGAGGAAT 1837  
QY 1501 GATCTCTCCACATTTGGTCCAGAGACTTTTGTGTGATGAAGAGGAAGTCTCTGGCTCGTGAT 1560

Db	1838	GAATCTCTCCACATGGTGAAGACTTCTTTGTGATGAAGAGAAAGTCTCGCTCGTAT	1897
Qy	1561	TTTGACCTGGGGTCTCAGGAATTCAGAAAGTGTAGTAATGCATGCCATACAGCTGCTG	1620
Db	1898	TTTGACCTGGGGTCTCAGGAATTCAGAAAGTGTAGTAATGCATGCCATACAGCTGCTG	1957
Qy	1621	GGAATTTGTGTACAATACCCACCACTAGCAGGAACGATGATTTTTTATCACCCCCAGC	1680
Db	1958	GGAATTTGTGTACAATACCCACCACTAGCAGGAACGATGATTTTTTATCACCCCCAGC	2017
Qy	1681	ACAACCTGCCATCAGTCTTCGAATCACTTCTACAGCAATGGGGTACTTATGCTTTT	1740
Db	2018	ACAACCTGCCATCAGTCTTCGAATCACTTCTACAGCAATGGGGTACTTATGCTTTT	2077
Qy	1741	ATCATGGAGGCCATCATAGCTTGACGCTTTATGCACTTCTGAACAAGAGGGACTGGGG	1800
Db	2078	ATCATGGAGGCCATCATAGCTTGACGCTTTATGCACTTCTGAACAAGAGGGACTGGGG	2137
Qy	1801	GGTCCCACTAGCACCCCACTAACCTGATCAGCCAGGAGCAGCTGGTCCGGAAGCGGCC	1860
Db	2138	GGTCCCACTAGCACCCCACTAACCTGATCAGCCAGGAGCAGCTGGTCCGGAAGCGGCC	2197
Qy	1861	AGCCTGTCTACCTTCTCTCCAATCAAGGCACCATCTCAGTCCCTTGCAGACATTTTAC	1920
Db	2198	AGCCTGTCTACCTTCTCTCCAATCAAGGCACCATCTCAGTCCCTTGCAGACATTTTAC	2257
Qy	1921	CAAGTCTGCCATGAACAGTAGTAAGATTTATCCAGTAGTGGCATTTCTACAGTGGCAG	1980
Db	2258	CAAGTCTGCCATGAACAGTAGTAAGATTTATCCAGTAGTGGCATTTCTACAGTGGCAG	2317
Qy	1981	CACGATGACCCAGGAAGATATCAGTCTCTAGTCTTGTCTGAGCAGCAGTGGGCAAGAAGCTT	2040
Db	2318	CACGATGACCCAGGAAGATATCAGTCTCTAGTCTTGTCTGAGCAGCAGTGGGCAAGAAGCTT	2377
Qy	2041	CCAGAACCTTTGTCTTGGAAGATGATGAAGAATGAAGACAGTGACTTTGGGGAGGAA	2100
Db	2378	CCTGAACCTTTGTCTTGGAAGATGATGAAGAATGAAGACAGTGACTTTGGGGAGGAA	2437
Qy	2101	CAGCAGAGATTGCTACTCGAAGTGAGCCAATCCAGAGGACCAACAGCTTTTATCACCTTC	2160
Db	2438	CAGCAGAGATTGCTACTCGAAGTGAGCCAATCCAGAGGACCAACAGCTTTTATCACCTTC	2497
Qy	2161	TTACAGAGACTCTTGGGCGCTTGTCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC	2220
Db	2498	TTACAGAGACTCTTGGGCGCTTGTCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC	2557
Qy	2221	AACCTTCACTGCTGTTTCCAGAACCTCAGTATCTGCAAAAGTTGCAAAATACCTAATA	2280
Db	2558	AACCTTCACTGCTGTTTCCAGAACCTCAGTATCTGCAAAAGTTGCAAAATACCTAATA	2617
Qy	2281	ACCAGAACAGAAAGAAATGTTGCAGTATATCTGAGAGTGCCACATATTTCTTGTGAAG	2340
Db	2618	ACCAGAACAGAAAGAAATGTTGCAGTATATCTGAGAGTGCCACATATTTCTTGTGAAG	2677
Qy	2341	AATGCTGTCAAAATGTTTAAAGATATTTGGGGTTTTCAGAGGACCAAAACAAAGAGAGTG	2400
Db	2678	AATGCTGTCAAAATGTTTAAAGATATTTGGGGTTTTCAGAGGACCAAAACAAAGAGAGTG	2737
Qy	2401	TCTGTTTTAGAACTGAGCAGCACTTTTCTAATCAATGCAACCGCAAAAACTTCTAGAA	2460
Db	2738	TCTGTTTTAGAACTGAGCAGCACTTTTCTAATCAATGCAACCGCAAAAACTTCTAGAA	2797
Qy	2461	TATATTTCTGAGTTTGTGCTGCTG 2484	
Db	2798	TATATTTCTGAGTTTGTGCTGCTG 2821	

RESULT 5  
BC030783  
LOCUS Homo sapiens, KIAA1560 protein, clone MGC:26846 IMAGE:4821342, mRNA, complete cds.  
DEFINITION

ACCESSION BC030783  
VERSION BC030783.1  
KEYWORDS GI:21315057  
SOURCE MGC.  
ORGANISM Homo sapiens (human)  
REFERENCE 1 (bases 1 to 3490)  
AUTHORS Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (31-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbe-remail.nih.gov](mailto:cgapbe-remail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 33 Row: b Column: 19  
This clone was selected for full length sequencing because it passed the following selection criteria: GenomesCan gene prediction, Similarity but not identity to protein.

FEATURES  
Location/Qualifiers  
1..3490

/organism="Homo sapiens"

/mol\_type="mrna"

/db\_xref="LocusID:57678"

/db\_xref="taxon:9606"

/clone="MGC:26846 IMAGE:4821342"

/tissue\_type="Brain, hippocampus"

/clone\_lib="NIH MGC\_95"

/lab\_host="DH10B"

/note="Vector: pBluescript"

179..2665

/codon\_start=1

/product="KIAA1560 protein"

/protein\_id="AAH30783.1"

/db\_xref="GI:21315058"

/translation="MDEYALTLTIDVSLPHSSEYSGRCKHTSEEGCGFRPTIF

RSATLWKESIMSRKRPVGRCCVSCYCTQSQWDKFNTPISLGLRNVIYINETHRHR  
GWLARLSYVLFTQERDVHKGMPATNTENVLSSRQVEAIAEVAELNPDGSAQQQS  
KAVNKVKKAKRILQEMVATVSPAMIRLTGWLVLKLFNSFFWNIQIHKGQLEMVKAAT  
ETNLPLFLPVHRSHIDYLLLTFLFCHNKAPYASGNLNINIDIFSLIKLGGFFI  
RRDDETDRKQVLYRALLHGVILLRQQQFLIEFLEGRSRSGKTSKARAGLLSV  
VVDLTNVPIDILIIPVGISYDRIIEGHVNGEOLGPKKNESLWSVARGVIRMLRKN  
YGCYVDPAQPSLSKYLESSQKPSVALLSLEQALLPAILPSPSDAADRGRTLSIN  
ESRNATESRRRLIANLAHEILFTASKSCAIMSTHVACLLIYRHRQGDISTLVED  
FFVWKEVLARDFDLGSNSDVVMHAIQLLGNCTVITHTSRNDEFITETSTVPSV  
FLNFYNGVLHVFIEMAIACSLYAVLNKRGGLGPTSTPNLISQQLVRAKASLCY  
LLSNEGTLSPCOTFYOVCHETVKGTYQYGLITVAEHDDQDISPSLAEQMDKLPE  
PLWRSDEEDSDFGEOEDCYLKVSKSQHQOQFITFLORLLGLEAYSSAIFVH  
NFGSPVPEPEYLOKLHKLITRERNVAVYAESATYCLVNKNVQMKDVGFKETKQK  
RVSVLELSSTFLPQCNRQKLLVILSLFVL"

BASE COUNT 1002 a 734 c 790 g 964 t

ORIGIN

Query Match 99.8%; Score 2479.2; DB 9; Length 3490;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2481; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGATGAATCTGACCTGACCTGACCTGGTGGTACATAGATGTTCTTATCTGCCACATTCATCA 60  
|||||

179 ATGGATGAATATGCACTGACCCCTTGGTACAATAGATGTTTTCTTATCTGCCACATTCATCA 238  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

1141 GTAGCAAGAGGTGTTATTAGAAATGTTAGCAAAAAATATGTTGTTGTTCCGAGTGGATTTT 1200  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000

QY 2221 AACTTCAGTGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCCACAAATACCTAATA 2280  
Db |||||  
QY 2399 AACTTCAGTGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCCACAAATACCTAATA 2458  
Db |||||  
QY 2281 ACCAGAACAGAAAGAAAGTTCCAGTATATCTGAGAGTGGCAGATATCTGTTGTGAAG 2340  
Db |||||  
QY 2459 ACCAGAACAGAAAGAAAGTTGTCAGTATATCTGAGAGTGGCAGATATCTGTTGTGAAG 2518  
Db |||||  
QY 2341 AATGCTGTGAAATGTTTAAAGGATATGGGGTTCCTCAAGAGGACCAACAAAGAGAGCTG 2400  
Db |||||  
QY 2519 AATGCTGTGAAATGTTTAAAGGATATGGGGTTCCTCAAGAGGACCAACAAAGAGAGCTG 2578  
Db |||||  
QY 2401 TCTGTTTAACTGAGCAGCACTTTTCTACTCAATGCAACCGCACAAAACCTCTAGAA 2460  
Db |||||  
QY 2579 TCTGTTTAACTGAGCAGCACTTTTCTACTCAATGCAACCGCACAAAACCTCTAGAA 2638  
Db |||||  
QY 2461 TATATCTGAGTTTGTGCTGCTG 2484  
Db |||||  
QY 2639 TATATCTGAGTTTGTGCTGCTG 2662  
Db |||||

RESULT 6  
HSM804404 4918 bp mRNA linear PRI 30-APR-2003  
DEFINITION Homo sapiens mRNA; cDNA DKFZp451P0819 (from clone DKFZp451P0819);  
complete cds.  
ACCESSION AL833093  
VERSION AL833093.2 GI:30268382  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 4918)  
AUTHORS Bloeker, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C.,  
Fobo, G., Han, M., Osanger, A. and Wiemann, S.  
TITLE Direct Submission  
JOURNAL Submitted (30-APR-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANY  
COMMENT On Apr 30, 2003 this sequence version replaced gi:21733684.  
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
This clone (DKFZp451P0819) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
information about the clone and the sequencing project is available  
at http://mips.gsf.de/proj/cDNA/  
FEATURES  
Location/Qualifiers  
1..4918  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="10q25.2"  
/clone="DKFZp451P0819"  
/tissue\_type="cDNA-collection"  
/clone\_lib="451 (synonym: hicci). Vector pSport1; host  
DH10B; sites NotI + SalI"  
/dev\_stage="adult"  
1..4918  
/gene="DKFZp451P0819"  
200..2686  
/gene="DKFZp451P0819"  
/note="glycerol-3-phosphate acyltransferase"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAD8932.1"  
/db\_xref="GI:30268383"  
/translation="MDESALTGLGIDVSLPHSSEYSVGRCKHTSEWGECCFRPTIF  
RSATLKWKESLMRKPRGRCYSCSPQSWDKFNPISLGLRNVIIYNETHRHR  
GWLARLSYVLFIQERDVHKGWATNTVENLNSRVQEAIAEVAALNPDGSAQQQS

KAVNKVKAKRILQEMVATVSPAMIRLTGVLKLFNSFWNFQIHKGQLEMVKAAT  
ETNLLEFLVHRSHYDILLTFLFCHNIKAPYIASGNLNIPIFSTLHKGLGFFI  
RRLNDETPDKRQVLYRALLHGHVTELRQOFLIEI FLEGRSSGKTSACARAGLSV  
VVDTLSTNVIPIDILIPVGISYDRI EGHVNGEOLGPKKNESIWSVARGVIRLMKN  
YGCVRDPAOPFSLKEYLESQSKPVSALLSLEOALLPAILPSPSADREGRTLSIN  
ESRNATDESRRRLIANLASHIPTASCALIMSTHIVACILLYRHRGGIDLSLTVS  
FFVMKEVLARDFDLGFSGNSDVVMAIQLLGNCTVITHTSRNDEFFITPSTTVPSV  
FELNFGVGLVHVFIMEAI IACSLYAVLNKRGSGPTSTPNLISQBLVRKAASLCY  
LLSNEGTSILPCTOTFYOVCHETKGFIRYGLITVAEHDDOEDI SPSLAEOQWDKLP  
PLSWRSDEEDSDPGEORQDYLKVSQSEHOQFITFLORLLGLLEANSVAALFVH  
NFGSPVEPYLQKLHKYLITRERNVAVAESATYCLVKNVAMFKDIGNVFKETQK  
RVSVLEUSITFLPQCNKQKULEYILSFVVL"  
4893  
polyA\_site  
BASE COUNT 1379 a 1005 c 1075 g 1459 t  
ORIGIN  
Query Match 99.8%; Score 2479.2; DB 9; Length 4918;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2481; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 ATGGATGAATCTGCACCTGACCCCTTGGTACATAGATGTTTCTTATCTGCCACATTCATCA 60  
Db |||||  
QY 200 ATGGATGAATCTGCACCTGACCCCTTGGTACATAGATGTTTCTTATCTGCCACATTCATCA 259  
Db |||||  
QY 61 GAATACAGATGTTGTCGATGTAAGCACACAAGTCAGGAATGGGTGAGTGCGCTTTAGA 120  
Db |||||  
QY 260 GAATACAGATGTTGTCGATGTAAGCACACAAGTCAGGAATGGGTGAGTGCGCTTTAGA 319  
Db |||||  
QY 121 CCCACCTCTTCAGATCTGCAACTTTAAATGGAAAGAACGCCCTTAATGAGTCGGAAGAG 180  
Db |||||  
QY 320 CCCACCTCTTCAGATCTGCAACTTTAAATGGAAAGAACGCCCTTAATGAGTCGGAAGAG 379  
Db |||||  
QY 181 CCATTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAATTTTCAAC 240  
Db |||||  
QY 380 CCATTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAATTTTCAAC 439  
Db |||||  
QY 241 CCCAGTATCCCGTCTTTGGGTTTCCGGAATGTTATTTATATCAATGAAATCACAACAAGA 300  
Db |||||  
QY 440 CCCAGTATCCCGTCTTTGGGTTTCCGGAATGTTATTTATATCAATGAAATCACAACAAGA 499  
Db |||||  
QY 301 CACCGCGGATGGCTTCAAGACGCTTTCTTACCGTTCTTTTATTTCAAGACGCGAGATGTG 360  
Db |||||  
QY 500 CACCGCGGATGGCTTCAAGACGCTTTCTTACCGTTCTTTTATTTCAAGACGCGAGATGTG 559  
Db |||||  
QY 361 CATGAAGGCGATGTTTGGCCCAATGTCAGTGAATAATGCTGAACAGAGTAGAGTACAA 420  
Db |||||  
QY 560 CATGAAGGCGATGTTTGGCCCAATGTCAGTGAATAATGCTGAACAGAGTAGAGTACAA 619  
Db |||||  
QY 421 GAGGCAATTGCAAGAGTGGCTGCTGAATTAACCCCTGATGTTCTGCCACGACCAATCA 480  
Db |||||  
QY 620 GAGGCAATTGCAAGAGTGGCTGCTGAATTAACCCCTGATGTTCTGCCACGACCAATCA 679  
Db |||||  
QY 481 AAAGCCGTTAAACAAGTGAAGAAAGAAAGAGGATTTCTTCAAGAAATGGTTGCCACT 540  
Db |||||  
QY 680 AAAGCCGTTAAACAAGTGAAGAAAGAAAGAGGATTTCTTCAAGAAATGGTTGCCACT 739  
Db |||||  
QY 541 GTCTACCGGCAATGATCAGCTGACCTGGGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db |||||  
QY 740 GTCTACCGGCAATGATCAGCTGACCTGGGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 799  
Db |||||  
QY 601 TTTTGGAAACATTCAAATTCACAAAGTCAACTTCAGATGTTTAAAGTGGCAACTGAGACG 660  
Db |||||  
QY 800 TTTTGGAAACATTCAAATTCACAAAGTCAACTTCAGATGTTTAAAGTGGCAACTGAGACG 859  
Db |||||  
QY 661 AATTGCGCGTCTCTGTTTCTACAGATTCATAGATCCCATATTTGACTATCTGCTGCTCACT 720  
Db |||||  
QY 860 AATTGCGCGTCTCTGTTTCTACAGATTCATAGATCCCATATTTGACTATCTGCTGCTCACT 919  
Db |||||  
QY 721 TTCATTCCTTCTGCCATAACATCAAAAGCACCACATATTTGCTTTCAGGCAATAATCTCAAC 780  
Db |||||  
QY 920 TTCATTCCTTCTGCCATAACATCAAAAGCACCACATATTTGCTTTCAGGCAATAATCTCAAC 979  
Db |||||

```
QY 781 ATCCCAATCTTCAGTACCTTTGATCCATAAGCTTTGGGGCTTCTTTTCATACGACGAAGGCTC 840
Db |||
QY 980 ATCCCAATCTTCAGTACCTTTGATCCATAAGCTTTGGGGCTTCTTTTCATACGACGAAGGCTC 1039
Db |||
QY 841 GATGAACACACGATGGAACGGAAGAGTGTCTCTATAGAGCTTTGTCTCCATGGGCATATA 900
Db |||
QY 1040 GATGAACACACGATGGAACGGAAGAGTGTCTCTATAGAGCTTTGTCTCCATGGGCATATA 1099
QY |||
QY 901 GTTGAATTAATTCGACGACGACCAATCTTCGAGATCTTCCTGGAAGGCAACGCTTCTAGG 960
Db |||
QY 1100 GTTGAATTAATTCGACGACGACCAATCTTCGAGATCTTCCTGGAAGGCAACGCTTCTAGG 1159
QY |||
QY 961 AGTGGAAAAACCTCTTGTGTCTCGGGCAGAGACTTTTGTCACTGTGTGGTAGATACCTGTCT 1020
Db |||
QY 1160 AGTGGAAAAACCTCTTGTGTCTCGGGCAGAGACTTTTGTCACTGTGTGGTAGATACCTGTCT 1219
QY |||
QY 1021 ACCAATGTCTATCCAGACATCTTTGATTAATACCTGTGTGGAATCTCTATGATCCATATC 1080
Db |||
QY 1220 ACCAATGTCTATCCAGACATCTTTGATTAATACCTGTGTGGAATCTCTATGATCCATATC 1279
QY |||
QY 1081 GAAGGTCACTACATGCTGAACAACTGGGCAAACTTAAGAGAAATGAGAGCCTGTGGAGT 1140
Db |||
QY 1280 GAAGGTCACTACATGCTGAACAACTGGGCAAACTTAAGAGAAATGAGAGCCTGTGGAGT 1339
QY |||
QY 1141 GTAGCAAGAGGTGTATTAGAAATGTTACGAAAAAACTATGGTGTGTCCGAGTGGATTTT 1200
Db |||
QY 1340 GTAGCAAGAGGTGTATTAGAAATGTTACGAAAAAACTATGGTGTGTCCGAGTGGATTTT 1399
QY |||
QY 1201 GCACAGCATTTTCTTTAAAGGAATATTAGAAAGGCAAAAGTCAAGAAACCGGTGTCTGCT 1260
Db |||
QY 1400 GCACAGCATTTTCTTTAAAGGAATATTAGAAAGGCAAAAGTCAAGAAACCGGTGTCTGCT 1459
QY |||
QY 1261 CTACTTTCCCTGGAGCAGAGCTTTGTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320
Db |||
QY 1460 CTACTTTCCCTGGAGCAGAGCTTTGTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1519
QY |||
QY 1321 GCTGATGAAGGTGAGACACGCTCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380
Db |||
QY 1520 GCTGATGAAGGTGAGACACGCTCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1579
QY |||
QY 1381 CGAAGGAGGTGTGATTCGAAATCTGGGTGAGCATATTTCTATTCTCTAGCAAGTCTGT 1440
Db |||
QY 1580 CGAAGGAGGTGTGATTCGAAATCTGGGTGAGCATATTTCTATTCTCTAGCAAGTCTGT 1639
QY |||
QY 1441 GCCATTATGTCACACACATTTGGCTTGGCTGCTCTCTACAGACACAGGAGGGAAT 1500
Db |||
QY 1640 GCCATTATGTCACACACATTTGGCTTGGCTGCTCTCTACAGACACAGGAGGGAAT 1699
QY |||
QY 1501 GATCTCTCCACATTTGGTCGAAGACTCTTTGTGATGAAGAGGAAGTCCCTGGCTCGTGAT 1560
Db |||
QY 1700 GATCTCTCCACATTTGGTCGAAGACTCTTTGTGATGAAGAGGAAGTCCCTGGCTCGTGAT 1759
QY |||
QY 1561 TTTGACCTGGGGTTCTCAGGAAATTCAGAAATGTAGTAATGATGCAATACAGCTGCTG 1620
Db |||
QY 1760 TTTGACCTGGGGTTCTCAGGAAATTCAGAAATGTAGTAATGATGCAATACAGCTGCTG 1819
QY |||
QY 1621 GGAATTTGTCTCAATATCACCACACTAGCAGGAACGATGAGTTTTTATCACCACCCAGC 1680
Db |||
QY 1820 GGAATTTGTCTCAATATCACCACACTAGCAGGAACGATGAGTTTTTATCACCACCCAGC 1879
QY |||
QY 1681 ACAACTGTCCCATCAGTCTTTCGAATCACTTCTACAGCAATGGGTACTTTCATGCTTTT 1740
Db |||
QY 1880 ACAACTGTCCCATCAGTCTTTCGAATCACTTCTACAGCAATGGGTACTTTCATGCTTTT 1939
QY |||
QY 1741 ATCATGAGGCCATCATAGCTTTCAGCTTTTATGCAAGTCTGAAACAGAGGGGACTGGGG 1800
Db |||
QY 1940 ATCATGAGGCCATCATAGCTTTCAGCTTTTATGCAAGTCTGAAACAGAGGGGACTGGGG 1999
QY |||
QY 1801 GGTCCCACTAGCAGCCCACTCACTGATCAGCCAGAGCAGCTGTGGAGAGGGGCC 1860
Db |||
QY 2000 GGTCCCACTAGCAGCCCACTCACTGATCAGCCAGAGCAGCTGTGGAGAGGGGCC 2059
QY |||
QY 1861 AGCTGTGCTACCTTCTCTCCAATGAAGGACCACCATCTCACTGCTTGGCCAGACATTTTAC 1920
```

```
Db |||
QY 2060 AGCTGTGCTACCTTCTCTCCAATGAAGGCCACCATCTCACTGCTTGGCCAGACATTTTAC 2119
QY |||
QY 1921 CAAAGTCGCCATGAACAGTAGAGAAAGTTTATCCAGTATGGCATTTTACAGTGGCAGAG 1980
Db |||
QY 2120 CAAAGTCGCCATGAACAGTAGAGAAAGTTTATCCGGTATGGCATTTCTTACAGTGGCAGAG 2179
QY |||
QY 1981 CACGATGACGACGAGAGATATCAGTCTCTAGTCTTGTCTGAGCAGCAGTGGGACAAGAGCTT 2040
Db |||
QY 2180 CACGATGACGACGAGAGATATCAGTCTCTAGTCTTGTCTGAGCAGCAGTGGGACAAGAGCTT 2239
QY |||
QY 2041 CCAGAACCTTTCTCTTGGGAAAGTGAAGAAAGATGAAGACAGTGAATTTGGGGAGGAA 2100
Db |||
QY 2240 CCAGAACCTTTCTCTTGGGAAAGTGAAGAAAGATGAAGACAGTGAATTTGGGGAGGAA 2299
QY |||
QY 2101 CAGCGAGATGCTACTCTGAGGTGAGCCCAATCCAGGAGCACCAGCAGTTTATCACCTTC 2160
Db |||
QY 2300 CAGCGAGATGCTACTCTGAGGTGAGCCCAATCCAGGAGCACCAGCAGTTTATCACCTTC 2359
QY |||
QY 2161 TTACAGAGACTCTTGGGGCTTTTGTGGAGGCTTACAGCTCTGTGCTGCTTTTGTTCAC 2220
Db |||
QY 2360 TTACAGAGACTCTTGGGGCTTTTGTGGAGGCTTACAGCTCTGTGCTGCTTTTGTTCAC 2419
QY |||
QY 2221 AACTTCAGTGTCTCTTTCAGAACCTGAGTATCTGCAAAAAGTTGCAAAAATACCTTAATA 2280
Db |||
QY 2420 AACTTCAGTGTCTCTTTCAGAACCTGAGTATCTGCAAAAAGTTGCAAAAATACCTTAATA 2479
QY |||
QY 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGGCCACATATTTGTCTGTGAAG 2340
Db |||
QY 2480 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGGCCACATATTTGTCTGTGAAG 2539
QY |||
QY 2341 AATGCTGTCAAAATGTTTAAAGGATATTTGGGTTTTTCAAGGAGACCAAAACAAAGAGAGTG 2400
Db |||
QY 2540 AATGCTGTCAAAATGTTTAAAGGATATTTGGGTTTTTCAAGGAGACCAAAACAAAGAGAGTG 2599
QY |||
QY 2401 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACCAAAAACCTTTAGAA 2460
Db |||
QY 2600 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACCAAAAACCTTTAGAA 2659
QY |||
QY 2461 TATATTCTGAGTTTTTGTGTGCTG 2484
Db |||
QY 2660 TATATTCTGAGTTTTTGTGTGCTG 2683
```

RESULT 7  
HSM804372

LOCUS HSM804372 6387 bp mRNA linear PRI 30-APR-2003  
DEFINITION Homo sapiens mRNA; cDNA DKFZp451B1115 (from clone DKFZp451B1115);  
complete cds.

ACCESSION AL833061  
VERSION AL833061.2 GI:30268346

## KEYWORDS

Source: Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 6387)  
Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C.,  
Fobo, G., Han, M., Osanger, A. and Wiemann, S.

## AUTHORS

Direct Submission

## TITLE

Submitted (30-APR-2003) MIPS, Ingolstaedter Landstr.1, D-85764

## JOURNAL

Neuberberg, GERMANY  
On Apr 30, 2003 this sequence version replaced gi:21733651.  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.

## COMMENT

This clone (DKFZp451B1115) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further  
information about the clone and the sequencing project is available  
at <http://mips.gsf.de/proj/cDNA/>.





```
QY 1501 GATCTCTCCACATTTGGTCGAAGACTTCTTTGTGATGAAAGAGAACTCTCTGGCTCGTGAT 1560
Db 1697 GATCTCTCCACATTTGGTCGAAGACTTCTTTGTGATGAAAGAGAACTCTCTGGCTCGTGAT 1756
QY 1561 TTTGACCTGGGTTTTCAGGAAATTCAGAAAGATGTAGTATGATGCCATACAGCTGCTG 1620
Db 1757 TTTGACCTGGGTTTTCAGGAAATTCAGAAAGATGTAGTATGATGCCATACAGCTGCTG 1816
QY 1621 GGAATTTGTGTACAAATCCACACTAGCAGGAAGATGAGTTTATATACCCCCAGC 1680
Db 1817 GGAATTTGTGTACAAATCCACACTAGCAGGAAGATGAGTTTATATACCCCCAGC 1876
QY 1681 ACAACTGTCCCATCAGTCTTCGAATCAACTTCTACAGCAATGGGTACTTCATGTCTTT 1740
Db 1877 ACAACTGTCCCATCAGTCTTCGAATCAACTTCTACAGCAATGGGTACTTCATGTCTTT 1936
QY 1741 ATCATGGAGGCCATCATAGCTTGCAGCCTTTATGCAATGTTCTGAAAGAGGGAAGCTGGGG 1800
Db 1937 ATCATGGAGGCCATCATAGCTTGCAGCCTTTATGCAATGTTCTGAAAGAGGGAAGCTGGGG 1996
QY 1801 GGTCCCACTAGCACCCACCTTAACCTGATCAGCCAGGACAGCTGTGGGAGGGCGCC 1860
Db 1997 GGTCCCACTAGCACCCACCTTAACCTGATCAGCCAGGACAGCTGTGGGAGGGCGCC 2056
QY 1861 AGCCTGTGCTACCTTCTCTCAATGAAGCACCATCTCACTGCTTGCAGACATTTTAC 1920
Db 2057 AGCCTGTGCTACCTTCTCTCAATGAAGCACCATCTCACTGCTTGCAGACATTTTAC 2116
QY 1921 CAACTGTGCATGAACACAGTAGGAAAGTTTATCCAGTATGGCAATCTTACAGTGGCAGAG 1980
Db 2117 CAACTGTGCATGAACACAGTAGGAAAGTTTATCCAGTATGGCAATCTTACAGTGGCAGAG 2176
QY 1981 CACGATGACCGAGAGATATCAGTCTAGCTTGTGCTGAGCAGCAGTGGGACAGAGCTT 2040
Db 2177 CACGATGACCGAGAGATATCAGTCTAGCTTGTGCTGAGCAGCAGTGGGACAGAGCTT 2236
QY 2041 CCAGAACCTTTGTCTTGGAGAGTGTAGAGAGATGAAGACAGTCACTTTTGGGAGGAA 2100
Db 2237 CTTGAACCTTTGTCTTGGAGAGTGTAGAGAGATGAAGACAGTCACTTTTGGGAGGAA 2296
QY 2101 CAGCGAGATTGCTACCTGAAGGTGAGCCAAATCCAGGAGACACAGAGTTTATCACCTTC 2160
Db 2297 CAGCGAGATTGCTACCTGAAGGTGAGCCAAATCCAGGAGACACAGAGTTTATCACCTTC 2356
QY 2161 TTACAGAGACTCTTGGGCTTTGTGGAGGCTTACAGCTCTGCTGCCATCTTGTGTCAC 2220
Db 2357 TTACAGAGACTCTTGGGCTTTGTGGAGGCTTACAGCTCTGCTGCCATCTTGTGTCAC 2416
QY 2221 AACTTCAGTGGTCTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTAATA 2280
Db 2417 AACTTCAGTGGTCTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTAATA 2476
QY 2281 ACCAGAACAGAAAGAAATGTGCAGTATATGCTGAGAGTGCCACATATTTGTCTGTGAAG 2340
Db 2477 ACCAGAACAGAAAGAAATGTGCAGTATATGCTGAGAGTGCCACATATTTGTCTGTGAAG 2536
QY 2341 AATGCTGTGAAATGTTTAAAGATATGTTGGGTTTTCAAGGAGCCAAACAAAGAGAGTG 2400
Db 2537 AATGCTGTGAAATGTTTAAAGATATGTTGGGTTTTCAAGGAGCCAAACAAAGAGAGTG 2596
QY 2401 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAACCTTCTAGAA 2460
Db 2597 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAACCTTCTAGAA 2656
QY 2461 TATATTCTCAGTTTTTGTGGTGTG 2484
Db 2657 TATATTCTCAGTTTTTGTGGTGTG 2680
```

## RESULT 8

HSMB03772  
LOCUS  
DEFINITION Homo sapiens mRNA; cDNA DKF2p313H071 (from clone DKF2p313H071).

ACCESSION AL832464  
VERSION AL832464.1 GI:21733030  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 5689)  
AUTHORS Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.  
TITLE Direct Submission  
JOURNAL Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY  
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.  
This clone (DKFZp313H071) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES  
Location/Qualifiers  
1..5689  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp313H071"  
/tissue\_type="human fetal skin"  
/clone\_lib="313 (synonym: hlcc2). Vector pTriplex2; host DH10B; sites SfiIA + SfiIB"  
/dev\_stage="fetal"  
5654  
polyA\_site  
BASE COUNT 1521 a 1149 c 1256 g 1763 t  
ORIGIN

Query Match 85.0%; Score 2111.8; DB 9; Length 5689;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2116; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 1 ATGGATGAATCTGCACTGACCCCTTGGTCAATAGATGTTTCTTATCTGCCACATTCATCA 60
Db 167 ATGGATGAATCTGCACTGACCCCTTGGTCAATAGATGTTTCTTATCTGCCACATTCATCA 226
QY 61 GAATACAGTGTGGTTCGATGTAAGCACACAAAGTGAAGATGGGGTGAAGTGGCTTTAGA 120
Db 227 GAATACAGTGTGGTTCGATGTAAGCACACAAAGTGAAGATGGGGTGAAGTGGCTTTAGA 286
QY 121 CCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCTAATGAGTCGGAAGG 180
Db 287 CCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCTAATGAGTCGGAAGG 346
QY 181 CCATTTGTTGGAAGATGTTGTTACTCTCTGCACTCCCGAGAGCTGGGACAAAATTTTCAAC 240
Db 347 CCATTTGTTGGAAGATGTTGTTACTCTCTGCACTCCCGAGAGCTGGGACAAAATTTTCAAC 406
QY 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAACTCACACAAGA 300
Db 407 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAACTCACACAAGA 466
QY 301 CACCGGGATGGCTTGCAGAGCCCTTTCTTACGTTCTTTTATTTTCAAGAGCAGATG 360
Db 467 CACCGGGATGGCTTGCAGAGCCCTTTCTTACGTTCTTTTATTTTCAAGAGCAGATG 526
QY 361 CATAAGGGCATGTTTGGCCACCAATGTGACTGAAAAATGCTGAAACAGCAGTAGAGTACAA 420
Db 527 CATAAGGGCATGTTTGGCCACCAATGTGACTGAAAAATGCTGAAACAGCAGTAGAGTACAA 586
QY 421 GAGGCAATTCAGAGATGGCTGCTGAAATTAACCCCTGATGGTTCTGCCAGAGCAATCA 480
Db 587 GAGGCAATTCAGAGATGGCTGCTGAAATTAACCCCTGATGGTTCTGCCAGAGCAATCA 646
```



QY	481	AAAGCGGTAAACAAGTCAAAAAGAAAGACTTAAAGGATTCCTTCAAGAAATGGTTGCCACT	540
DB	647	AAAGCGGTAAACAAGTCAAAAAGAAAGACTTAAAGGATTCCTTCAAGAAATGGTTGCCACT	706
QY	541	GTCTCACCGGCAATGATCAGACTGACATGGGTGGGTGCTGCTTAAAACTGTTTCAACAGCTTC	600
DB	707	GTCTCACCGGCAATGATCAGACTGACATGGGTGGGTGCTGCTTAAAACTGTTTCAACAGCTTC	766
QY	601	TTTTGGAAATTTCAAAATTCACAAAGGTCACCTTGAGATGGTTAAAGCTGCAACTGAGACG	660
DB	767	TTTTGGAAATTTCAAAATTCACAAAGGTCACCTTGAGATGGTTAAAGCTGCAACTGAGACG	826
QY	661	AATTTGCCGCTTCTGTTTCTACCAAGTTCATAGATCCCATATTTGACTATCTGCTGCTCACT	720
DB	827	AATTTGCCGCTTCTGTTTCTACCAAGTTCATAGATCCCATATTTGACTATCTGCTGCTCACT	886
QY	721	TTCAATTTCTTTCTGCCATAACATCAAAAGCACCATACATTTGCTTCAGGCAATAATCTCAAC	780
DB	887	TTCAATTTCTTTCTGCCATAACATCAAAAGCACCATACATTTGCTTCAGGCAATAATCTCAAC	946
QY	781	ATCCCAATCTTCAGTACCTTGATCGATTAAGCTTGGGGCTTCTTCATACAGACGAAGCTC	840
DB	947	ATCCCAATCTTCAGTACCTTGATCGATTAAGCTTGGGGCTTCTTCATACAGACGAAGCTC	1006
QY	841	GATGAAACACCAAGATGGACGAAAGATGTTCTCTATAGAGCTTTGCTTCATGGGCATATA	900
DB	1007	GATGAAACACCAAGATGGACGAAAGATGTTCTCTATAGAGCTTTGCTTCATGGGCATATA	1066
QY	901	GTTGAATTTACTTCGACAGCAGCAATCTTTGGAGATCTTCTCGAAGGCACACGTTCTTAGG	960
DB	1067	GTTGAATTTACTTCGACAGCAGCAATCTTTGGAGATCTTCTCGAAGGCACACGTTCTTAGG	1126
QY	961	AGTGGAAAAAAGCTCTTTGTGCTCGGCGAGGACTTTTGTTCAGTTGTGGTAGATACTCTGTCT	1020
DB	1127	AGTGGAAAAAAGCTCTTTGTGCTCGGCGAGGACTTTTGTTCAGTTGTGGTAGATACTCTGTCT	1186
QY	1021	ACCAATGTCATCCAGACATCTTGATTAATACCTGTGGAATCTCTATGATCCGATTC	1080
DB	1187	ACCAATGTCATCCAGACATCTTGATTAATACCTGTGGAATCTCTATGATCCGATTC	1246
QY	1081	GAAGGTCACTACAATGGTGAACAACCTGGGCAAACTTAAGAAAGATGAGAGCTCTGGAGT	1140
DB	1247	GAAGGTCACTACAATGGTGAACAACCTGGGCAAACTTAAGAAAGATGAGAGCTCTGGAGT	1306
QY	1141	GTAGCAAGAGGTGTTATTAGAAATGTTACGAAAAAATCTATGGTTGTGTCGAGTGGATTTT	1200
DB	1307	GTAGCAAGAGGTGTTATTAGAAATGTTACGAAAAAATCTATGGTTGTGTCGAGTGGATTTT	1366
QY	1201	GCACAGCAATTTTCTTTAAAGGAATATTTAGAAAGCCAAAGTCCAGAAACCGGTGCTGCT	1260
DB	1367	GCACAGCAATTTTCTTTAAAGGAATATTTAGAAAGCCAAAGTCCAGAAACCGGTGCTGCT	1426
QY	1261	CTACTTTCCCTGGAGCAAGCGTTGTTACCAAGCTATATCTCTTCAAGACCCAGTGAATGCT	1320
DB	1427	CTACTTTCCCTGGAGCAAGCGTTGTTACCAAGCTATATCTCTTCAAGACCCAGTGAATGCT	1486
QY	1321	GCTGATGAAGGTAGAGACAGTCCATTAATGAGTCCGAAATGCAACAGATGAATCCCTTA	1380
DB	1487	GCTGATGAAGGTAGAGACAGTCCATTAATGAGTCCGAAATGCAACAGATGAATCCCTTA	1546
QY	1381	CGAAGGAGTTGATTTGCAAAATCTGGCTGAGCATATCTATTCTACTGCTAGCAAGTCTGT	1440
DB	1547	CGAAGGAGTTGATTTGCAAAATCTGGCTGAGCATATCTATTCTACTGCTAGCAAGTCTGT	1606
QY	1441	GCCATTAATGTCCACACATTTGTGGCTTGCTGCTCTCTACAGACACAGCGAGGAAAT	1500
DB	1607	GCCATTAATGTCCACACATTTGTGGCTTGCTGCTCTCTACAGACACAGCGAGGAAAT	1666
QY	1501	GATCTCTCCCATTTGGTTCGAGACTTCTTTGTGATGAAGAGGAGTCTCTGGCTCTGTAT	1560
DB	1667	GATCTCTCCCATTTGGTTCGAGACTTCTTTGTGATGAAGAGGAGTCTCTGGCTCTGTAT	1726

QY	1561	TTTGACCTGGGGTTCTCAGGAAATTCAGAGATGTAGTAATGCATGCCATACAGCTGCTG	1620
DB	1727	TTTGACCTGGGGTTCTCAGGAAATTCAGAGATGTAGTAATGCATGCCATACAGCTGCTG	1786
QY	1621	GGAAATGTGTGCACAAATCACCCACACTAGCAGGAACGATGAGTTTATTCACCCCCAGC	1680
DB	1787	GGAAATGTGTGCACAAATCACCCACACTAGCAGGAACGATGAGTTTATTCACCCCCAGC	1846
QY	1681	ACAACTGTCCCATCAGTCTTCGAACTCAACCTTACAGCAATGGGTACTTTCATGTCTTT	1740
DB	1847	ACAACTGTCCCATCAGTCTTCGAACTCAACCTTACAGCAATGGGTACTTTCATGTCTTT	1906
QY	1741	ATCATGGAGGCCATCATAGCTTTCAGCCTTTATCGATTCCTGAACAGAGGGGACTGGGG	1800
DB	1907	ATCATGGAGGCCATCATAGCTTTCAGCCTTTATCGATTCCTGAACAGAGGGGACTGGGG	1966
QY	1801	GGTCCCACTAGCACCCCACTTAACCTGATCAGCCAGGAGCAGCTGTGTCGGAAGCGGCC	1860
DB	1967	GGTCCCACTAGCACCCCACTTAACCTGATCAGCCAGGAGCAGCTGTGTCGGAAGCGGCC	2032
QY	1861	AGCCTGTGCTACTTCTCTCCAAATGAAGGCCACCACTCTCACTGCCCTTCCAGACATTTTAC	1920
DB	2027	AGCCTGTGCTACTTCTCTCCAAATGAAGGCCACCACTCTCACTGCCCTTCCAGACATTTTAC	2086
QY	1921	CAAGTCTGCCATGAACAGTAGTAGAAAGTTTATCCAGTATGGCATCTTACAGTGCAGAG	1980
DB	2087	CAAGTCTGCCATGAACAGTAGTAGAAAGTTTATCCAGTATGGCATCTTACAGTGCAGAG	2146
QY	1981	CACGATGACCAAGGAGATATCAGTCCCTAGTCTTCTGAGCAGCAGTGGGCAAGAAGCTT	2040
DB	2147	CACGATGACCAAGGAGATATCAGTCCCTAGTCTTCTGAGCAGCAGTGGGCAAGAAGCTT	2206
QY	2041	CCAGAACCTTTGCTTCTGGAGAAATGATGATGAAGAGATGAAGACAGTGTGTCGGGAGGAA	2100
DB	2207	CCTGAACTTTGCTTCTGGAGAAATGATGATGAAGAGATGAAGACAGTGTGTCGGGAGGAA	2266
QY	2101	CAGCGAGATGCTACTCTGAAGGT 2123	
DB	2267	CAGCGAGATGCTACTCTGAAGGT 2289	
RESULT 9			
AB046780		5674 bp mRNA linear PRI 22-FEB-2001	
LOCUS	AB046780	Homo sapiens mRNA for KIAA1560 protein, partial cds.	
DEFINITION	AB046780		
ACCESSION	AB046780.1	GI:10047184	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			

/clone="fhl1959"		/notes="vector:paluescriptII SK plus"		1159 AGAATGTTACGAAAAAATATGTTGTGTCGAGTGGATTTTGCACAGCAATTTTCTCTTA		1218
1. .5674		/genes="KIAA1560"		661 AGAATGTTACGAAAAAATATGTTGTGTCGAGTGGATTTTGCACAGCAATTTTCTCTTA		720
<1. .1989		/genes="KIAA1560"		1219 AAGGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCTCTACTTTCCCTCGAGAGCA		1278
/notes="Start codon is not identified."		/codon start=1		721 AAGGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCTCTACTTTCCCTCGAGAGCA		780
/product="KIAA1560 protein"		/protein_id="BAB13386.1"		1279 GGGTGTGTACCAAGCTATCTCTCTCAACACCCAGTCAGTGTCTGCTGATGAAGTGTAGAGAC		1338
/db_xref="GI:10047185"		/translation="KKAKRIQEMVATVSPAMIRLTGWVLLKLFNSFFWNQIHKQG		781 GGGTGTGTACCAAGCTATCTCTCTCAACACCCAGTCAGTGTCTGCTGATGAAGTGTAGAGAC		840
LEMKAAETENLPLELPHVRSHIDYLLTFLFCHNIRKAPYIASGNLNIPIFSLI		HKLGGFIRRLDETDPGRKDVLRLLHGHIVELLRQQOFLFIELEGRSRSGKTS		1339 AGCTCCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGGAGGTTGATTGCA		1398
ARAGLLSVVVDLSTNVPDILIPVIGISYDRIIEGHYNGEOLGKPKNESLWSVARG		VIRMLKNYGRVDPAPQSLKEYLESOSQKPVSAALLSLEQALLPAILPSRSDAAD		841 AGCTCCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGGAGGTTGATTGCA		900
EGDTSINERSNATDESRLRLJANLAHLLPTASKSCALMSTHIVICLLLYRHRGDI		DLSTLVEDFVMEELARDIDFGSNESEDVVMHAIQLLGNCTVTHLTSRNDFFIT		1399 AATCTGGCTGAGCATATTTCTATTCTAGTAGAAGTCTCTGTGCCATTTATGTCACACAC		1458
PSTTVSFVFNLSNGVLHVIMEAIIACSLYVNLKRLGGFTSPNLIISQEQLV		RKAASLYLLSNGETISLPQTFYQVCHETGVKFIQYGLITVAEHDDQDIIPSLAEQ		901 AATCTGGCTGAGCATATTTCTATTCTAGTAGAAGTCTCTGTGCCATTTATGTCACACAC		960
QWDKLPPELSWRSDDEDSDFGEQRDCYLKVSQKEHQOFTIFLQRLGLPLEAY		SSAAIIVHNFSGVPPEVYLQKLHKLITRTERNVAVYAESATYCLVKNVXMFKDIG		1459 ATTGTGGCTTGGCTCTCTCTACAGACACAGCGAGGGAATTTGATCTCTCCACATTTGTC		1518
VFKEIKQKRVSLVLSSTFLPQCNQRKLLIYLISFVL"		BASE COUNT 1656 a 1087 c 1207 g 1724 t		961 ATTGTGGCTTGGCTCTCTCTACAGACACAGCGAGGGAATTTGATCTCTCCACATTTGGTC		1020
ORIGIN		Query Match 79.9%; Score 1984.4; DB 9; Length 5674;		1519 GAAGACTTCTTTGTGATGAAGAGGAAGTCTCTGGCTCTGCTGATTTTGACCTGGGGTTCTCA		1578
		Best Local Similarity 99.9%; Pred. No. 0;		1021 GAAGACTTCTTTGTGATGAAGAGGAAGTCTCTGGCTCTGCTGATTTTGACCTGGGGTTCTCA		1080
		Matches 1985; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		1579 GGAATTTCAAGAGATGTAGTAATGCAATGCAATACAGCTGCTGGGAAATTTGTGTCAAAATC		1638
QY 499 AAAAAAGAGCTAAAGGATTTCTCAAGAAATGGTTGCCACTCTCTCACCGGCAATGATC		558		1081 GGAATTTCAAGAGATGTAGTAATGCAATACAGCTGCTGGGAAATTTGTGTCAAAATC		1140
DB 1 AAAAAAGAGCTAAAGGATTTCTCAAGAAATGGTTGCCACTCTCTCACCGGCAATGATC		60		1639 ACCCACTAGCAGAGAACGATGAGTCTTTTATCAACCCAGCAGCACTCTCCATCAGTC		1698
QY 559 AGACTGACTGGGTGGTCTGCTAAACGTTTCAACAGCTCTCTTTGGAACTTTCAAAAT		618		1141 ACCCACTAGCAGAGAACGATGAGTCTTTTATCAACCCAGCAGCACTCTCCATCAGTC		1200
DB 61 AGACTGACTGGGTGGTCTGCTAAACGTTTCAACAGCTCTCTTTGGAACTTTCAAAAT		120		1699 TTTGCAACTCAACTTTTACAGCAATGGGTGATCTTCAATGTTATGATGAGGCACTATA		1758
QY 619 CACAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACGAATTTGCGCTTCTGTTT		678		1201 TTTGCAACTCAACTTTTACAGCAATGGGTGATCTTCAATGTTATGATGAGGCACTATA		1260
DB 121 CACAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACGAATTTGCGCTTCTGTTT		180		1759 GCTTGAGCCTTTATGCAAGTCTGAAACAGAGGAGTGGGGGGTCCCACTAGCACCCCA		1818
QY 679 CTACCAAGTTCATAGATCCCATATTTGACTATCTGCTGCTCACTTTCAATCTCTTCTGCCAT		738		1261 GCTTGAGCCTTTATGCAAGTCTGAAACAGAGGAGTGGGGGGTCCCACTAGCACCCCA		1320
DB 181 CTACCAAGTTCATAGATCCCATATTTGACTATCTGCTGCTCACTTTCAATCTCTTCTGCCAT		240		1819 CCTAACCTGATCAGCCAGGAGCAGCTGGTGCGGAAGGCGGCGGCTGTGCTACCTTCTC		1878
QY 739 AACATCAAGACCATACATTTGCTTCAGGCAATTAATCTCAACATCCCAATCTTCAGTACC		798		1321 CCTAACCTGATCAGCCAGGAGCAGCTGGTGCGGAAGGCGGCGGCTGTGCTACCTTCTC		1380
DB 241 AACATCAAGACCATACATTTGCTTCAGGCAATTAATCTCAACATCCCAATCTTCAGTACC		300		1879 TCCAATGAAGGCAACCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAAACA		1938
QY 799 TTGATCCATGAAGCTTGGGGCTTTCTTATACGACGAAGGCTCGATGAACACACAGATGGA		858		1381 TCCAATGAAGGCAACCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAAACA		1440
DB 301 TTGATCCATGAAGCTTGGGGCTTTCTTATACGACGAAGGCTCGATGAACACACAGATGGA		360		1939 GTAGGAAGTTTATCCAGTATGCGCATTTTACAGTGGCAGACAGATGACAGGAAGAT		1998
QY 859 CGAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATAGTTGAAATTTACTTCGACAG		918		1441 GTAGGAAGTTTATCCAGTATGCGCATTTTACAGTGGCAGACAGATGACAGGAAGAT		1500
DB 361 CGAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATAGTTGAAATTTACTTCGACAG		420		1999 ATCAGTCTCAGTCTTGTGCTGAGCAGTGGGACAGAGCTTTCCAGAACCTTTGCTTGG		2058
QY 919 CAGCAATCTTGAGATCTCTCGGAAGGCACAGCTTCTAGGAGTGGAAAAACCTCTTGT		978		1501 ATCAGTCTCAGTCTTGTGCTGAGCAGTGGGACAGAGCTTTCCAGAACCTTTGCTTGG		1560
DB 421 CAGCAATCTTGAGATCTCTCGGAAGGCACAGCTTCTAGGAGTGGAAAAACCTCTTGT		480		2059 AGAAGTGTATGAAGAAGATGAAGAAGTGTCTTTGGGAGGAAACAGCGAGTTGTACCTG		2118
QY 979 GCTCGGGCAGGACTTTTGTGCTGAGTGTGGTAGATCTGCTACCAATGTCAATCCAGAC		1038		1561 AGAAGTGTATGAAGAAGATGAAGAAGTGTCTTTGGGAGGAAACAGCGAGTTGTACCTG		1620
DB 481 GCTCGGGCAGGACTTTTGTGCTGAGTGTGGTAGATCTGCTACCAATGTCAATCCAGAC		540		2119 AAGGTGAGCCAAATCOAAGGAGCACAGCAGTTTATCACCTTTTACAGAGATCTCTTGGG		2178
QY 1039 ATCTTGATAATACCTGTGGAATCTCCTATGATCGCAATTTATGAAGGTCACTCAATGTT		1098		1621 AAGGTGAGCCAAATCOAAGGAGCACAGCAGTTTATCACCTTTTACAGAGATCTCTTGGG		1680
DB 541 ATCTTGATAATACCTGTGGAATCTCCTATGATCGCAATTTATGAAGGTCACTCAATGTT		600		2179 CTTTTGCTGGAGCCTACAGCTCTGCTGCATCTTTGTTTGTTCACAACTTCAAGTGTCTGTT		2238
QY 1099 GAACAACTGGGCAAACTTAAGAAAGATAGAGGCTGTGGAGTGTAGCAAGAGTGTATT		1158		1681 CCTTTGCTGGAGCCTACAGCTCTGCTGCATCTTTGTTTGTTCACAACTTCAAGTGTCTGTT		1740
DB 601 GAACTGCTGGGCAAACTTAAGAAAGATAGAGGCTGTGGAGTGTAGCAAGAGTGTATT		660				

QY 2239 CCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTAATAACAGAACAGAAAGAAAT 2298  
Db 1741 CCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTAATAACAGAACAGAAAGAAAT 1800  
QY 2299 GTTGCAATATGCTGAGAGTGCACATATGCTGTTGGAAGATGCTGTTGGAAGATGTTT 2358  
Db 1801 GTTGCAATATGCTGAGAGTGCACATATGCTGTTGGAAGATGCTGTTGGAAGATGTTT 1860  
QY 2359 AAGGATATGAGGTTTCAAGGAGACCAAAACAAAGAGAGTCTGTTTGAAGACTGAGC 2418  
Db 1861 AAGGATATGAGGTTTCAAGGAGACCAAAACAAAGAGAGTCTGTTTGAAGACTGAGC 1920  
QY 2419 AGCACTTTTACCTCAATGCAACCGACAAACAACTTTAGAAATATATCTGAGTTTGTG 2478  
Db 1921 AGCACTTTTACCTCAATGCAACCGACAAACAACTTTAGAAATATATCTGAGTTTGTG 1980  
QY 2479 GTGCTG 2484  
Db 1981 GTGCTG 1986

RESULT 10  
AX401972  
LOCUS AX401972 2646 bp DNA linear PAT 06-JUN-2002  
DEFINITION Sequence 1648 from Patent WO0210453.  
ACCESSION AX401972  
VERSION AX401972.1 GI:21338152  
KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1  
AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and  
Elashoff, M.R.  
TITLE Molecular toxicology modeling  
JOURNAL Patent: WO 0210453-A 1648 07-FEB-2002;  
Gene Logic, Inc. (US)  
FEATURES  
source  
1..2646  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="caxon:10116"  
/note="EMBL/GenBank Accession No. NM\_017274"

BASE COUNT 670 a 689 c 670 g 617 t  
ORIGIN

Query Match 75.7%; Score 1879.2; DB 6; Length 2646;  
Best Local Similarity 84.8%; Pred. No. 0;  
Matches 2106; Conservative 0; Mismatches 378; Indels 0; Gaps 0;

QY 1 ATGATGATCTGCACTGACCTGTTGTTCAATAGATGTTTCTATCTGCCACATTCATCA 60  
Db 160 ATGGAGGAGTCTTCAGTGACAAATTTGGCACAATAGACGTTTCTATCTGCCAAATCATCA 219  
QY 61 GAATACAGTGTGTCGATGTAAGCACACAGTACGGAATCGGCTGAGTGGCTTTAGA 120  
Db 220 GAATACAGCTTGGCCGATGTAACACACAGTACGGAATCGGCTGAGTGGCTTTCAA 279  
QY 121 CCCACCGTCTTCAGATCTGCAATTTTAAATGGAAGAAAGCCTAATGAGTCGGAAGAG 180  
Db 280 CTTACCTTTTCAGATCCGCAACGCTGAAATGGAAGAGAGCCTCATGAGCCGGAAGAG 339  
QY 181 CCAATTTGTTGAAGATGTTGTTTACCTGCACTCCCAAGAGCTGGGCAAAATTTTCAAC 240  
Db 340 CCTTCTGTTGGGAAGTGTGCTATTCATGACGCTCAGAGCTGGGAAGGTTTTCAC 399  
QY 241 CCCAGTATCCCGTCTTTGGGTTGGGATCTTATTTATATCAATGAATCAACACAGA 300  
Db 400 CCCAGTATCCCATCTCTGGGTTGGGAATGTTATTTATATCAATGAATCAACACAGG 459  
QY 301 CACCGCGATGGCTTGCAAGACGCTTCTTACGTTCTTTTATTTCAAGACGAGATGTG 360

Db 460 CACCGAGATGGCTGGCAAGACGCTTCTTACATCCTTTTGTTCAGAGCGCATGTC 519  
QY 361 CATAAGGGCATGTTTGGCCCAATGTGACCTGAAATCTGTGACACAGCTAGAGTACAA 420  
Db 520 CACAGGGCATGTTTGGCCCAATGTGACCTGAAATCTGTGACACAGCTAGAGTACAA 579  
QY 421 GAGCAATTCAGAAAGTGGCTGTGAATTAACCTGATGTTCTGCCACAGCAGCAATCA 480  
Db 580 GAGCAATTCAGAAAGTGGCTGTGAATTAACCTGATGTTCTGCCACAGCAGCTCC 639  
QY 481 AAGCCCTTTAACAAGTGAAGAAAGAAAGCTTAAAGGATTTTCAAGAAATGTTTGCAC 540  
Db 640 AAGCCATTCAGAAAGTGAAGAAAGAAAGCTTAAAGGATTTTCAAGAAATGTTTGCAC 699  
QY 541 GTCTCACGGCAATGATCAGACTGACTGGGTGGGTGCTGTAAACTGTTTCAACAGCTTC 600  
Db 700 GTCTCACGGCAATGATCAGACTGACTGGGTGGGTGTTACTAAAGCTTCTCAACAGCTTC 759  
QY 601 TTTTGAACATTTCAAAATTCAGAAAGTCAACTTCAGATGTTTAAAGCTGCAACTGAGACG 660  
Db 760 TTTTGAACATTTCAAAATTCAGAAAGTCAACTTCAGATGTTTAAAGCTGCAACTGAGACG 819  
QY 661 AATTGCGCTTCTGTTTCTACAGTTTCAATAGATCCCATATTTGATATCTGCTCACT 720  
Db 820 AATCTGCGCTTCTGTTTCTGCGGTGCAAGATCCCATGACTACCTGCTGCTCACT 879  
QY 721 TTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
Db 880 TTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939  
QY 781 ATCCCAATTTTCAAGTACCTTTGATCCTAAGCTTGGGGGCTTCTTATACAGCAAGGCTC 840  
Db 940 ATCCCAATTTTCAAGTACCTTTGATCCTAAGCTTGGGGGCTTCTTATACAGCAAGGCTC 999  
QY 841 GATGAAACACAGATGCAAGCAAGATGTTTCTCTATAGAGCTTTGCTTCCATGGGCATATA 900  
Db 1000 GATGAAACACAGATGCAAGCAAGATGTTTCTCTATAGAGCTTTGCTTCCATGGGCATATA 1059  
QY 901 GTTGAATTTCTGACAGAGCAATTTCTTGAGATCTTCTGAGATCTTCTGGAAGGCAAGTCTAGG 960  
Db 1060 GTTGAATTTCTGACAGAGCAATTTCTTGAGATCTTCTGAGATCTTCTGGAAGGCAAGTCTAGG 1119  
QY 961 AGTGAAGAAACCTCTTCTGCTCGGGCAGGACTTTTGTGAGTGTGTTGTTGTTGTTGTTGTT 1020  
Db 1120 AGTGAAGAAACCTCTTCTGCTCGGGCAGGACTTTTGTGAGTGTGTTGTTGTTGTTGTTGTT 1179  
QY 1021 ACCAATGTCATCCAGACATCTTTGATTAATACCTTGTGGAATCTCTATGATCGCATATC 1080  
Db 1180 TCCAACACCATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCT 1239  
QY 1081 GAAGGTCACTACATGTTGTAACAACTGGGCAAACTTGAAGAAATGAGAGCTGTGAGT 1140  
Db 1240 GAAGGTCACTACATGTTGTAACAACTGGGCAAACTTGAAGAAATGAGAGCTGTGAGT 1299  
QY 1141 GTAGCAAGGTGTTTATGATGTTTACGAAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200  
Db 1300 GTAGCAAGGTGTTTATGATGTTTACGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1359  
QY 1201 GCACAGCATTTTCTTAAAGGAATATTTAAGAAAGCCAAAGTCAAGAAACCGGCTCTGCT 1260  
Db 1360 GCACAGCATTTTCTTAAAGGAATATTTAAGAAAGCCAAAGTCAAGAAACCGGCTCTGCT 1419  
QY 1261 CTACTTTTCCCTGGAGCAAGGTTGTTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320  
Db 1420 CTTCTCTTTTGGAGCAAGCACTGTTTACAGCAATCTTCTTCAAGACCTGTGCTGCT 1479  
QY 1321 GCTGATGAAGGTAGAGACAGTCCATTAATGAGTCCAGAAATGCAACAGATGATCCCTA 1380  
Db 1480 GCTGCGCAACATGAAGACATGTTCCAGTAATGAGTCCAGAAATGCAACAGATGATCCCTA 1539  
QY 1381 CGAAGGAGGTGTTGTTGCAAAATCTGGCTGAGCATATTTCTTCTGCTGCTGCTGCTGCTGCT 1440

Db 1540 CGAAGGAGGCTGATCGCAAAACCTGGGGAGACATTTCTTACCAGCCAGAGTCTCTGC 1599  
QY 1441 GCCATTATGTCACACACATTTGGCTGCTGCTCTCTACAGACACAGGAGGGAATT 1500  
Db 1600 GCTATCATGTCCACCATTTGGCTGCTGCTCTCTACAGACACAGGAGGGAATC 1659  
QY 1501 GATCTCTCCACATTTGGTGAAGACTTTCTTTGTGATGAAGAGGAAGTCTCTGCTGAT 1560  
Db 1660 CACCTCTCCACGCTGGTGAAGACTTTCTTTGTGATGAAGAGGAAGTCTAGTCCGGAT 1719  
QY 1561 TTTGACCTGGGTTCTCAGAAATTCAGAAATGTAGTATGATGATGATGATGATGATGATG 1620  
Db 1720 TTTGACCTGGGTTCTCCGGAATTCAGAAATGTAGTATGATGATGATGATGATGATGATG 1779  
QY 1621 GGAATTTGTCTCAATCACTCCACACTAGCAGAAAGATGATGATGATGATGATGATGATG 1680  
Db 1780 GGAATTTGTCTCAATCACTCCACACTAGCAGAAAGATGATGATGATGATGATGATGATG 1839  
QY 1681 ACAACTGTCCCATCAGTCTTCGAACCTTCTACAGCAATGGGTACTTTCATGCTTT 1740  
Db 1840 ACAACTGTCCCGTCCGCTTTGAACCTTCTACAGCAATGGGTACTTTCATGCTTT 1899  
QY 1741 ATCATGAGGCAATCATAGTCTGAGCTTTATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 1800  
Db 1900 ATCATGAGGCAATCATAGTCTGAGCTTTATGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 1959  
QY 1801 GGTCCCACTAGCACCCACTTACCTGATCAGCAGGAGCAGTGTGGGAGGGCGCC 1860  
Db 1960 GGTCTGCGGAGGCTTGGCAACCTGATCAGCAGGAGCAGTGTGGGAGGGCGCC 2019  
QY 1861 AGCTGTGCTACCTTCTCTCAATGAAGCAGCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 1920  
Db 2020 AGCTGTGCTACCTTCTCTCAATGAAGTACATTTCTGCTGCTGAGTCTGAGTCTGAGTCTG 2079  
QY 1921 CAAGTCTGCCATGAACAGTAGGAAGTTTATCAGTATGAGTCTTACAGTGGCAGAG 1980  
Db 2080 CAGGTTGTCAAGAGACAGTAGGAAGTTTATCAGTATGAGTCTTACAGTGGCAGAG 2139  
QY 1981 CAGGATGACAGGAGATACAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 2040  
Db 2140 CAGGATGACAGGAGATACAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 2199  
QY 2041 CCAGAACTTTGTCTTGGAGAGTGTAGGAAGATGAAGACAGTCTGAGTCTGAGTCTGAGTCTG 2100  
Db 2200 CCGAGGCTCTGAGTGGAGAGTGTAGGAAGATGAAGACAGTCTGAGTCTGAGTCTGAGTCTG 2259  
QY 2101 CAGGAGATGCTTACCTGAAGTGTAGGAGTGTAGGAAGATGAAGACAGTCTGAGTCTGAGTCTG 2160  
Db 2260 CAGGATGATGCTTACCTGAAGTGTAGGAGTGTAGGAAGATGAAGACAGTCTGAGTCTGAGTCTG 2319  
QY 2161 TTACAGAGTCTTGGGCTTTGCTGGAGGCTTACAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 2220  
Db 2320 CTGACAGGCTTCTGGGCTTTGCTGGAGGCTTACAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 2379  
QY 2221 AACTCTGAGTCTGCTGTTCCAGAACTGAGTATCTGCAAAAGTGTGCAAAATACCTAATA 2280  
Db 2380 ACCTTCCGCGCCAGTCCCGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 2439  
QY 2281 ACCAGAACAGAAAGAAATGTTGAGTATGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2340  
Db 2440 ACCAGAACAGAAAGAAATGTTGAGTATGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2499  
QY 2341 AATGCTGTGAATGTTTGAAGTATTTGGGTTTCAAGAGACCAACAAAGAGAGTGT 2400  
Db 2500 AATGCTGTGAATGTTTGAAGTATTTGGGTTTCAAGAGACCAACAAAGAGAGTGT 2559  
QY 2401 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAACCTTTCTAGAA 2460  
Db 2560 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAACCTTTCTAGAA 2619  
QY 2461 TATATTCTGAGTTTGTGGTGTGCTG 2484  
Db 2620 TACATTCTGAGCTTGTGGTGTGCTG 2643

## RESULT 11

AF021348

## LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .2646

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/tissue\_type="liver"

1. .2646

/gene="GPAT"

160. .2646

/gene="GPAT"

/EC\_numbers="2.3.1.15"

/note="similar to protein encoded by Mus musculus glycerol

3-phosphate acyltransferase gene, GenBank Accession Number

W7003, and Accession Numbers U36771, U36772, U36773"

/codon\_start=1

/product="sn-glycerol-3-phosphate acyltransferase"

/protein\_id="A4B71605.1"

/db\_xref="GI:2444459"

/translation="MESSVTITGTDVSLPNSBSYSLGRCKHTNEDWDCGKPTFF

RSATLKWESLMSRKRPFCYSCYCTPTDNLNLSRVOEIAEVAELNPNDSAQOOS

GWLARLVLVQERDVHGMFATSDNLNLSRVOEIAEVAELNPNDSAQOOS

KAIQKVKARKILOEMVATVSPGMRLTGWVLLKLPNSFFWNIQIHKGQLEWKAAT

ETNLELFLVHRSHIDLLTLFLFLCHNIKAPYASNNLNIPFSTLHLKLGFFI

RRRLDTPGRKDILYRALLHGHIVELLRQQQFLFLEIGTRSRSGKTSACARGLLSV

VDTLSNTIPDILVPGISYDRIIEGHNGEQLGPKKNESLMSVARGVIRMLRN

YGVYVDFAQFSLKEYLEGQSKVSPALSLQALLPAILPSRDPDAAAEHEDSSN

BSRNAADAFRRRLIANLAHELITASKSCAIMSTHVACLLLYRHRQIHLSTLVED

PEVMEELVARDPDLGFSNSEDVVMHAIQLLGNCTVITHTSRKDEFFTPSTTVPSV

FELNYSNGVLVFMIEALISYAVQNRKSGSGSAGGLNLISQELVRAASLCY

LUSNGTSLPCQTFYVQCETVGRPIQIGLTVAAQDDQEDVPSQLAEPQKKLIFE

PLNWSDEDESDSFEQRQCYLQSKAQHQQTFLQRLGLLPLLEAYSSAAIFVH

TPRGVPSEYILQHLRYLLTRERNVAVYAESATYCLVKNVAKMFKDGVFKETKQK

RASVLELSTFLPQSGRQKLELYILSFVVL"

BASE COUNT 670 a 689 c 670 g 617 t

ORIGIN

Query Match

Best Local Similarity 75.7%; Score 1879.2; DB 10; Length 2646;

Matches 2106; Conservative 0; Mismatches 378; Indels 0; Gaps 0;

QY 1 ATGGATGAATCTGCACTGACCTTGGTCAATAGATGTTCTTATCTGCCATTCATCA 60

Db 160 ATGGAGGAGTCTTCAGTGAATGGCAATAGACGTTTCTTATCTGCCAATTCATCA 219

QY 61 GAATACAGTGTGGTCGATGTAAGCACACAAGTGAAGAAATGGGGTGAAGTGGCTTTAGA 120  
Db 220 GAATACAGCTTTGGCCGATGTAAACACACAGATGAGGACTGGGTGACTGTGGCTTCAAA 279  
QY 121 CCCACCGCTCTCAGATCTGCAACTTTTAAATGGAAGAAAGCCTAATCAGTCGGAAGAG 180  
Db 280 CCTACTCTTTCAGATCCGCAACGCTGAATGGAAGGAGAGCCTCATGAGCCGGAAGAGG 339  
QY 181 CCAATTGTTGGAAGATGTTGTTACTCTTGCACTCCCCAGAGCTGGGACAAATTTTTCAAC 240  
Db 340 CCCCTCGTGGGAAGGTGTGCTATTATCATGCAGCCTCAGAGCTGGGAAGGTTTTTCAAC 399  
QY 241 CCCAGTATCCGCTCTTTGGGTTGGGAATGTTTATTAATCAATGAATCACTACACAAGA 300  
Db 400 CCCAGTATCCCATCTCTGGGTTTGGGAATGTTTATTAATCAATGAATCACTACACAAGG 459  
QY 301 CACC CGGATGCTTGCAGAGCGCTCTTTCTTACGTTCTTTTATTCAAGAGCGAGATGTG 360  
Db 460 CACCAGGATGGCTGGAAGACGGCTTTCTTACATCCCTTTTGTTCGAAGCGCGATGTC 519  
QY 361 CATAAGGCAATGTTTGCACCAATGTGACTGAAATGTGCTGAAACAGCAGTAGAGTACAA 420  
Db 520 CACAAGGCAATGTTTGCACCACTAGTACTGACATGTACTGAATAGCAGCAGTCCAA 579  
QY 421 GAGGCAATTCGAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCGCAGCGCAATCA 480  
Db 580 GAGGCAATTCGTAAGTGGCTGCAAAATTGAACCGGATGGATCTGCCCAGCAGTCC 639  
QY 481 AAAGCCGTTAAACAAGTGAAGAAAGAACTAAAGGATCTTCAAGAAATGGTTGCCACT 540  
Db 640 AAAGCCATCCAGAAGTGAAGAAAGAAAGCCAGGAAGATCCCTCCAGGAATGGTTGCTACA 699  
QY 541 GTCTCACCGGCAATGATCAGACTGAGTGGGTGGTCTGCTAAACTGTTCAAACAGCTTC 600  
Db 700 GTCTCCCCGGGATGATCAGGCTGACTGGCTGGGTGTTACTAAAGCTCTTCAACAGCTTC 759  
QY 601 TTTTGGAAATTCAAATTCAAAGGTCAAACTTGAGATGGTTAAAGCTGCAACTGAGACG 660  
Db 760 TTCTGGAAATTCAGATTCAAAGGTCAAACTTGAGATGGTGAAGCTGCAACTGAGACG 819  
QY 661 AATTGGCGCTCTGTTTCTACCACTTATAGATCCCATATGACTACTGCTGCTACT 720  
Db 820 AATCTGCCGCTCTGTTTCTGCGGTGACAGATCCCATCGACTACTGCTGCTACC 879  
QY 721 TTCAATCTCTTGTGCATAAATCAAGCAACATATGCTTTCAGGCAATAATCTCAAC 780  
Db 880 TTCAATCTCTTGTGCATAAATCAAGCTCCATACATCGCTCGGCAACAACTCAAC 939  
QY 781 ATCCCAATCTTTCAGTACCTTGATCCATAGCTTGGGGCTTCTTCATACGACGAGGCTC 840  
Db 940 ATCCCAATCTTTCAGTACCTTGATTCACAAAGCTTGGGGCTTTCATGAAGACGAGGCTT 999  
QY 841 GATGAACACAGATGGAGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCAATATA 900  
Db 1000 GACGAAACTCCAGATGGAAGCAACATCTGTGAAGAGGCTTCTCATGGGCATATA 1059  
QY 901 GTTGAATTTCTGACAGCAATCTTGTGAGATCTTCTGGAAGCTTCTGGAAGGACACGTTCTAGG 960  
Db 1060 GTTGAATTTCTGACAGCAATCTTGTGAGATCTTCTGGAAGGACACGCTTCCGC 1119  
QY 961 AGTGAAGAAACCTCTTGTGCTCGGCGAGGACTTTTGTGAGTGTGTTAGTACTCTGTCT 1020  
Db 1120 AGTGAAGAAACCTCTTGTGCTCGGCGAGGCTTCTGTCAGTGGTAGTACTCTGTCA 1179  
QY 1021 ACCAATGTATCCCAAGACATCTTGAATATACTGTTGAATCTCTCTATGATCGCATATC 1080  
Db 1180 TCCAACCACTCTGACATCTGCTCATCTGCTGAGGCTCTCTGATGATCGGATATC 1239  
QY 1081 GAAGTCTACTCAATGGTGAACACTGGGCAACCTTAAGAAATGAGAGCTGTGGAGT 1140  
Db 1240 GAAGTCTACTCAATGGTGAACACTGGGCAAGCTTGGGCAAGCTTGAAGTCTCTGGAGT 1299  
QY 1141 GTAGCAAGAGGTGTTATTAGAATGTTACGAAAAAACTATGTTGTTGTCGAGTGGATTTT 1200

Db 1300 GTGCAAGAGCGTTTATCAGAATGTCGGAAGAACTACCGCTATGTCCGAGTGGACTTT 1359  
QY 1201 GCACAGCATTTTCTTTAAAGGAATATTTAAAGCCAAAGCTCAGAAACCGTGTCTGCT 1260  
Db 1360 GCACAGCATTTTCTTTAAAGGAATATTTAAAGCCAAAGCTCAGAAACCTGTATCTGCT 1419  
QY 1361 CTACTTTCCCTGGAGCAAGCGTTGTACCACTATATCTTCAAGACCCAGTGAATGCT 1320  
Db 1420 CCCCTCTTTTGGAGCAAGCACTGTATCCAGCAATCTTCTTCAAGACCTGTATGCTGCT 1479  
QY 1321 GCTGATGAGGTGAGACACGTCCTCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db 1480 GCTCCGAACATTAAGACATGTCCAGTAATGAGTCGAGAAACGGGCGACAGCAAGCCCTC 1539  
QY 1381 CGAAGGAGGTGATTTGCAAAATCTGCGCTGAGCATATCTTATTCAGTGTAGCAAGTCTGCT 1440  
Db 1540 CGAAGGAGGTGATTCGCAAACTGGCGAGCACATCTCTTCCCGCAGCAAGTCTCTGC 1599  
QY 1441 GCCATTATGTCCACACACATTTGTGGCTTGGCTGCTCTCTACAGACACAGGCAAGAAAT 1500  
Db 1600 GCTATCATGTCCACCCACATTTGTGCTGCTGCTCTCTACAGACACAGGCAAGATC 1659  
QY 1501 GATCTCTCCATTTGTGCGAAGACTTCTTTGTGATGAAAGAGAGTCTCTGGCTGCTGAT 1560  
Db 1660 CACCTCTCCACGCTGGTGAAGACTTCTTTGTGATGAAGAGAGAGTCTCTAGCTCGGAT 1719  
QY 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAATGTAAGTATGTAATGCAATGCAATCAGTGTG 1620  
Db 1720 TTTGACCTGGGCTTCTCGGGAATTCAGAAATGTAAGTATGTAATGCAATCAGTGTG 1779  
QY 1621 GGAATTTGTGTCACAAATCACCCACACTAGCAGGAACATGATGATTTTTTATCACCCCAAGC 1680  
Db 1780 GGGAACTGTGTCACAAATCACCCACACTAGCAGGAAGATGATTTTTTATCTCTCCAGC 1839  
QY 1681 ACAACTGTCCCATCAGTCTTTCGAACTCAACTTTCAGCAATGAGGCTACTTCACTGTCTTT 1740  
Db 1840 ACAACTGTCCGTCGCTTTCGAACTCAACTTTCAGCAATGAGGCTACTTCACTGTCTTT 1899  
QY 1741 ATCATGAGGCGCATCATAGCTTGCAGCTTTATGCACTTCTGAACTGTAACAGAGGGGACTGGG 1800  
Db 1900 ATCATGAGGCGCATCATAGCTTGCAGCTTTATGCACTTCTGAACTGTAACAGAGGGGTTCCGGA 1959  
QY 1801 GGTCCCACTAGCACCCCACTTAACCTGATCAGCAGCAGCAGCTGTTGCGAAAGCGGCC 1860  
Db 1960 GGGTCTCGGAGGCTTGGCACTGATCAGCCAGAGAGAGCTGTTGCGAAAGCGGCC 2019  
QY 1861 AGCTGTGCTACTCTCTCCAATGAAGGCACTTCTCACTGCTCTGCAAGCACTTTTAC 1920  
Db 2020 AGCTGTGCTACTCTCTCTAATGAAGGTACCAATTTCTCTGCTGCTGCAAGCACTTTTAC 2079  
QY 1921 CAAGTCTGCCATGAAACAGTAGGAAGTTTATCAGTATGGCATTTCTTACAGTGGCAGAG 1980  
Db 2080 CAGGTTTGTCAAGAGACAGTAGGAAGTTTATCAGTACGGAATTTCTCACAGTGGCAGAG 2139  
QY 1981 CACCATGACGAGGAAGATATCAGTCTGCTGAGCAGCAGCTGTTGGAAGCAAGCTT 2040  
Db 2140 CAAGTACGAGGAAGATATCAGTCTGCTGAGCAGCAGCTGTTGGAAGCAAGCTT 2199  
QY 2041 CCAGAACTTTGTCTTGGAGAAAGTGAAGAAAGTGAAGCACTGTAAGTGGGAGGAA 2100  
Db 2200 CCGAGGCTCTGAACCTGGAGAAAGTGAAGAAAGTGAAGCACTGTAAGTGGGAGAG 2259  
QY 2101 CAGGAGATTTGCTTACTGAAGGTGAGCAATTCGAAGGAGCAACAGAGTATTTATCACCTTC 2160  
Db 2260 CAGGCTGATTTGCTTACTGAAGGTGAGCAGGCCAAGGAGCAACAGCAATTTATCACCTTT 2319  
QY 2161 TTACAGAGATCTCTTGGGCTTTGCTGAGGCTTACAGCTCTGCTGCACTGTTGTTAC 2220  
Db 2320 CTGAGAGGCTTCTGGGGCTTCTGAGGCTTACAGCTCTGCTGCACTTCTGCTGCACT 2379  
QY 2221 AACTTTAGTGTCTGTTTTCAGAACTGAGTATCTGCAAAAGTTCGCAAAATACCTAATA 2280











Db	2367	GCTATCATGTCCACCCACATTTGTGGCCTGTCTCTCTTACAGACACAGCGAGGGAATC	2425
Qy	1501	GATCTCTCCAAITTTGGTCGAAGACTTCTTTTGTATGAAGAGGAAGTCCCTGGCTCGTGAT	1560
Db	2427	CATCTCTCCAGCTTGTGGAAAGACTTCTTTTGTATGAAGGAGGAAGTCTTAGCTCGCGAT	2486
Qy	1561	TTTGACCTGGGGTTCTCAGGAAATTCAGAGATGTAGTAAATGCAATGCATACAGCTCGTG	1620
Db	2487	TTTGACCTTAGGCTTCTCCGGAAATTCAGAAAGATGTCTGTCATGCAATGCTATTTCAGCTTCTG	2546
Qy	1621	GGAAATTTGTGCACAAATCACCCACACTAGCAGGAAACGATGAGTTTATTTATCACCCCCAGC	1680
Db	2547	GGGAACTGTGCACAAATCACCCACAGCAGGAGGAAGATGAGTTTATTTATCTCCGACG	2606
Qy	1681	ACAACTGTCCCATCAGTCTTCGAACTCAACTTTACAGCAATAGGGGTACTTCATGTCTTT	1740
Db	2607	ACAACTGTCCGTCAGTCTTTTGAACCTTCTTACAGCAATGCGTACTTTCATGTGTTTC	2666
Qy	1741	ATCATGAGGCGCATCATAGCTTTCAGAGCTTTATGSCAGTTCTGAACAGAGGGGACTGGGG	1800
Db	2667	ATCATGGAAGCCATCATAGCTTTCAGACATCTATGSCAGTCTTGAATAAGAGGTGCTCTGGA	2726
Qy	1801	GGTCCCACTAGCACCCACCTTAACCTTGATCAGCCAGGAGCAGCTGGTGCAGAAAGCGGCC	1860
Db	2727	GGGTCCGCTGGAGGCCCTCGGCAACCTTGATCAGCCAGGAGCAGCTGGTGAAGAGGCCGCC	2786
Qy	1861	AGCTGTGTACTCTTCTCCAATGAAGGACCAATCTCACTGCCTTGCCAGACATTTTAC	1920
Db	2787	AGCTGTGTACTCTTCTCTAAACGAAGTACCAATTTCTTGCCCTGCCAGACTTTTAC	2846
Qy	1921	CAAGTCTGCCATGAACACAGTAGGAAAGTTTATCCAGTATGGCATCTTTCAGATGGCAGAG	1980
Db	2847	CAAGTTTGTATGAGACAGTTTGGCAAGTTTCAATCCAGTATGGCATCTTCAGATGGCAGAG	2906
Qy	1981	CAGCATGACCAGGAAGATATCAGTCTTAGTCTTGTCTGAGCAGCAGTGGGCAAGAAGCTT	2040
Db	2907	CAAGATGACCAGGAAGATGTCAGTCTCTGGCCTTCAGAAACAGCAGTGGGCAAGAAGCTT	2966
Qy	2041	CCGAACCTTTGTCTTGAGAGATGATGAAGAGTGAAGAGTGAACACAGTGACTTTGGGAGGAA	2100
Db	2967	CCGGAAC---TGAACCTGGGAAGTGAAGGGAAGATGAACACAGTGACTTTGGTGAAGG	3023
Qy	2101	CAGCGAATTTGCTACCTGAAGGTGAGGCCAATCAAAGGAGCACACAGCAGTTTATCACCTTC	2160
Db	3024	CAGCGAATTTGCTATCTCAAGTGTAGCCAGTCCAAAGGAGCACACAGCAGTTTATCACCTTT	3083
Qy	2161	TTACAGAGACTCCTTTGGGCCCTTTGCTGGAGGCCCTTACAGCTCTGTCGCATCTTTGTTCAC	2220
Db	3084	CTTCAGAGGCTTTAGGTCCCTGCTAGAAGCCTTACAGCTCTGTCGCATCTTTGTTCAC	3143
Qy	2221	AACCTTCAGTGTCTCTGTTTCAGAACCTGAGTATCTGCAAAAGTTGCAACAATACCTTAATA	2280
Db	3144	AACTTCAGCGGTCCAGTTCGAGTCTGAGTACTCTGCAGAACTGCACAGTACCTTTATC	3203
Qy	2281	ACCAGAACAGAAAATAATTTGACAGTATATGCTGAGAGTGCCCATATTTGCTTTGTGAAG	2340
Db	3204	ACCAGGACGAAAGGAACGTTGCGGTGTAACGCTGAGAGTCCACATACCTCTCTCGTGAAG	3263
Qy	2341	AATGCTGTGAAATGTTTAGGATATTTGGGGTTTTTACAGAGAGCCAAACAAAAGAGAGTG	2400
Db	3264	AAACGCTGTGAAATGTTTAAAGACATCGGGTTTTTCAAAGAGAGCCAAAGCAAGCAGCGTG	3323
Qy	2401	TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAAAAATCTCTAGAA	2460
Db	3324	TCTGTTCTAGAACTGAGCAGTACTTTTCTTACCTCAGTGCACCGGAGAAATCTCTAGAG	3383
Qy	2461	TATATTTGAGTTTTTGTGGTGCTG	2484
Db	3384	TATATTTGAGTTTTTGTGGTGCTG	3407

RESULT 14  
RNU36771

LOCUS  
 DEFINITION RNU36771 2519 bp mRNA linear ROD 26-MAR-1999  
 Rattus norvegicus sn-glycerol 3-phosphate acyltransferase mRNA,  
 nuclear gene encoding mitochondrial protein, complete cds.  
 ACCESSION U36771 U36772 U36773  
 VERSION U36771.2 GI:4481754  
 KEYWORDS  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE  
 AUTHORS 1 (bases 1 to 2519)  
 TITLE Nikonov,A.V., Morimoto,T. and Haldar,D.  
 JOURNAL Properties, purification and cloning of mitochondrial sn-glycerol  
 3-phosphate acyltransferase  
 (in) Pandalai S.G. (Ed.),  
 RECENT RESEARCH DEVELOPMENTS IN LIPIDS RESEARCH, VOLUME 2: 207-222;  
 Transworld Research Network, India (1998)

REFERENCE  
 AUTHORS 2 (bases 1 to 2519)  
 TITLE Nikonov,A.V., Morimoto,T. and Haldar,D.  
 JOURNAL Direct Submission  
 Submitted (20-SEP-1995) Andrei V. Nikonov, Department of Biological  
 Science, St. John's University, Jamaica, NY 11439, USA

REFERENCE  
 AUTHORS 3 (bases 1 to 2519)  
 TITLE Nikonov,A.V., Morimoto,T. and Haldar,D.  
 JOURNAL Direct Submission  
 Submitted (23-MAR-1999) Cell Biology, New York University School of  
 Medicine, 550 1st Ave. MSB-697, New York, NY 10016, USA

REMARK  
 COMMENT Sequence update by submitter  
 On or before Mar 23, 1999 this sequence version replaced  
 gi:1754786, gi:1754788, gi:1754784.

FEATURES  
 source  
 1..2519  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="Sprague Dawley"  
 /db\_xref="taxon:10116"  
 /sex="male"  
 /tissue\_type="liver"  
 /clone\_lib="Stratage cat. number 936513"  
 101..2419  
 /BC\_number="2.3.1.15"  
 /note="similar to product encoded by Mus musculus glycerol  
 3-phosphate acyltransferase gene, GenBank Accession Number  
 M77003"  
 /codon\_start=1  
 /product="sn-glycerol 3-phosphate acyltransferase"  
 /protein\_id="AA339470.2"  
 /db\_xref="GI:4481755"  
 /translation="MSRKRPFGVRCYCSTPOSWERFFNPSPPLGLRNVIYNETH  
 RHRGLARLSYLVFOEDVHKGFATSTDNVLSRVQAEIAEIAELNDPDSQA  
 QSKALQKVKARKILQEMVAIVSGMIRLGLWLLKLFNSFFNIIHQGLEVMVR  
 ATEIRLDPFLFPHRSHYDILLHTFLFCHNIKAPYTAGNNLNIPISTLTHKLG  
 FIRRLDTPGRKDLRYALLHGVVELLRQQQFLFLEGTSRSGKTSKARAG  
 LVNVDVTSNTIPDILVPIGVISYDRQIEGHNGQELGPKKNESLMSVARGVIRML  
 RKNYGVRVDFQAPSLKEGQKPSAPLSLEOALLPAILPSRPDAAAHEHDM  
 SNSESNADEAFRELRILANAEHLLTASKCAIMSTHIVACLLYRHWGHLSTL  
 VDPTFMKEVLARDFLAFSGNSDEDVNHAIQLGNCVTITLTKRKHQILPSTIV  
 PSVFELVSYNGLVHVFINEALIASIYAKNGRGSGGLPEPGLISQBLVRKAASL  
 CYLLNEGTSILPCTQFTYGVHETGKFIYQIGLTVAEQDDQEDVSPGLAEQWKKL  
 PEPFLNRSDFDEEDDFGEQRDCYLKVSQAKEHQQFITFLORLLGPLLEAYSSAIF  
 VITFPAVPESYLOKLHRYLLTRTERNVAVTAESATYCLVKNAQMKFDIGVFKETK  
 QKRSVLELSTTFLLQSGRSQKLEVLSPVL"

CDS  
 transit\_peptide 101..286  
 /note="contains R(-10) motif"  
 misc\_feature 1367..1405  
 /note="transmembrane-region site"  
 /evidence=not\_experimental  
 misc\_feature 1658..1705  
 /note="transmembrane-region site"  
 /evidence=not\_experimental  
 misc\_feature 254..286  
 /note="encodes R(-10) motif recognizable by matrix

BASE COUNT		627 a		670 c		644 g		578 t		proteases"	
ORIGIN											
Query Match		72.6%;		Score 1803;		DB 10;		Length 2519;			
Best Local Similarity		84.4%;		Pred. No. 0;							
Matches 2041;		Conservative		0; Mismatches 375;		Indels 3;		Gaps 1;			
QY	66	CAGTGTGGTGCATGTAAGCACACACAGTGAAGGATGGGTGAGTGTGGCTTTAGAGCCAC	125								
DB	1	CAGCTTGGCGAGTGTAAACACAGAAAGAGGAGTGGCTTTGACTTTGGCTTTCAACCTAC	60								
QY	126	CGTCTTCAGATCTGCACACTTTAAATGGAAGAAAGCCTTAATGAGTCGGAAAGGCCATT	185								
DB	61	CTTCTTCAGATCCGCAACGCTGAAATGGAAGGAGAGCTCATGAGCGGGAAGAGGCCCTT	120								
QY	186	TGTTGGAGATGTTGTGTACTCTCGACTCCCGAGAGCTGGGACAAATTTTTTCAACCCGAG	245								
DB	121	CGTGGGAAGGTGTGTCTATTATGCAACGCTCAGAGCTGGGAAGGTTCTTCAACCCGAG	180								
QY	246	TATCCCGTCTTGGTGTGGGAATGTTATTATATCAATGAAGAACTCACACAGACACCG	305								
DB	181	TATCCCACTCTGGTGTGGGAATGTTATTATATCAATGAAGAACTCACACAGGACCG	240								
QY	306	CGGATGGCTTGCAGAGCGCTTCTTACAGTCTTTTTTATTCAAGAGCGAGATGTGCATTA	365								
DB	241	AGGATGGCTGCAAGAGCGCTTCTTACATCCTTTTGTTCAGAGCGCGATGTCCACAA	300								
QY	366	GGGATGTTTGGCACCAATGTGATGAAATGTCTGAAACAGCAGATAGATACAAAGGC	425								
DB	301	GGGATGTTTGGCACCAATGTGATGAAATGTCTGAAACAGCAGATAGATACAAAGGC	360								
QY	426	AATTGAGAGTGGCTGCTCAATTAACCCCTGATGTTCTGCCAGCAGCAATCAAAAGC	485								
DB	361	AATTGAGAGTGGCTGCTCAATTAACCCCGATGATCTGCCAGCAGCAATCAAAAGC	420								
QY	486	CGTTAAACAAAGTGAAGAAAGAGCTTAAAGAGTCTTCAAGAAATGGTTGCCACTGTCTC	545								
DB	421	CATCCAGAAAGTGAAGAAAGAGCCAGGAAGATCCTCCAGGAATGGTTGCTACAGTCTC	480								
QY	546	ACGGCAATGATCAGACTGACTGGGTGGGTGCTGTCTTAAACCTGTTCAACAGCTTCTTTG	605								
DB	481	CCCGGGATGATCAGGCTGACTGGCTGGGTGTTACTAAAGCTCTTCAACAGCTTCTTCTG	540								
QY	606	GAACATTCAAATTCACAAAGGTCACTTGAAGTGTAAAGCTGCACTGAGACGATTTT	665								
DB	541	GAACATTCAGATTCACAAAGGTCACTTGAAGTGTAAAGCTGCACTGAGACGATTTT	600								
QY	666	GCCTCTTCTGTTCTACAGTTCATAGATCCCATATTGACTATCTGCTGCTCACTTTTCAT	725								
DB	601	GCCTCTTCTGTTCTGCGGTGCACAGATCCACATCGACTACCTGTTGCTCACTTCAT	660								
QY	726	TCCTTCTGCGATTAACATCAAGACCATACATTTGCTTCAAGCAATTAATCTCAACATCCC	785								
DB	661	CCTCTTCTGCGATTAACATCAAGCTCCATACATCGCTCCGCGCAACCAACCTCAACATCCC	720								
QY	786	AATCTTCAGTACCTTGATCATAGCTTGGGGCTTCTTCATACGAGAGGCTCCGATGA	845								
DB	721	CATCTTCAGTACCTTGATCATAGCTTGGGGCTTCTTCATAGAGAGGCTTTCATAGAGGCTTTCACGA	780								
QY	846	AACACAGATGGAGGAAAGATGTTCTATAGAGCTTTTCTCCATGGGCATATAGTTGA	905								
DB	781	AATCCAGATGGAGGCAAGACATCTGTACAGAGCGTTGCTCCATGGGCATGATGTTGA	840								
QY	906	ATTACTTCGACAGAGCAATCTTGGAGATCTTCTGGAAAGGCAACAGTTCTAGGAGTGG	965								
DB	841	ACTCTCTCGACAGCAGCAGTCTCTGGAGATCTTCTGGAAAGGCAACCGCTCCGCGAGTGG	900								
QY	966	AAAAAATCTTGTCTGGGCGAGCTTTTGTGAGTGTGATGATCTCTGTCTACCA	1025								
DB	901	CAAGACCTCTGTCTGGGCGGCGCTCTGTGAGTGTGATGATCTCTGTCTATCCAA	960								
QY	1026	TGTCATCCACAGACATCTTGATAATACCTGTTGGAATCTCTATGATCGCATTCGAAGG	1085								
QY	961	CACCATCCCTGACATCCTGGTTCATCCCTGTGGGCATCTCGTATGATCGGATATCAAGG	1020								
QY	1086	TCACTACATGCTGACACTGGGAAACCTTAGAGAAATGAGCCTCTGGAGTGTAGC	1145								
DB	1021	TCACTACATGCTGACACTGGGAAACCTTAGAGAAATGAGCCTCTGGAGTGTAGC	1080								
QY	1146	AAGAGGTGTTATTAGAAATGTTACGAAATAATATGTTGTTGTCGAGTGGATTTTCACA	1205								
DB	1081	AAGAGGTGTTATTAGAAATGTTACGAAATAATATGTTGTTGTCGAGTGGATTTTCACA	1140								
QY	1206	GCATTTTCTTAAAGGAATATTTAGAAAGCAAAGTCAAGAACCGGTGTCTGTCTACT	1265								
DB	1141	GCATTTTCTTAAAGGAATATTTAGAAAGCAAAGTCAAGAACCGGTGTCTGTCTACT	1200								
QY	1266	TTCCCTGGAGCAAGCGTGTACCGACTATATCTCTTCAAGACCCAGTGTGCTCTGA	1325								
DB	1201	CTCTTTGGAGCAAGCACTGTTACCGAAATCTCTCTTCAAGACCTGTGCTGCTGC	1260								
QY	1326	TGAAGGTGAGACACGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAG	1385								
DB	1261	CGAAATGAGACATGTCAGTAATGAGTCGAGAAACGCGCAGACGAGCTTCCGAAG	1320								
QY	1386	GAGGTGATTTCAAAATCTGGCTGAGCATATCTATTTACTGTAGCAAGTCTCTGTGCCAT	1445								
DB	1321	GAGGTGATTTCAAAATCTGGCTGAGCATATCTATTTACTGTAGCAAGTCTCTGTGCCAT	1380								
QY	1446	TATGTCCACACATTTGTGGCTTGTCTCTCTCTACAGACAGCAGCGGAAATTCATCT	1505								
DB	1381	CATGTCCACCCCATTTGTGGCTTGTCTCTCTCTACAGACAGCAGCGGAAATTCACCT	1440								
QY	1506	CTCCACATTTGTGCAAGACTTCTTTGTGATGAAAGAGGAAGTCTTGGCTCGTGAATTTGA	1565								
DB	1441	CTCCACCTTGTGGAAGACTTCTTTGTGATGAAAGAGGAAGTCTTGGCTCGTGAATTTGA	1500								
QY	1566	CCTGGGGTCTCAGGAAATTCAGAAATGATAGTAATGATGCAATACAGTCTCTGGGAA	1625								
DB	1501	CCTGGGGTCTCAGGAAATTCAGAAATGATAGTAATGATGCAATACAGTCTCTGGGAA	1560								
QY	1626	TTGTGTCAATCAATCACCACACTAGCAGGAACGATGATTTTATCACCCCGACCAAC	1685								
DB	1561	CTGTGTCAATCAATCACCACACTAGCAGGAAGATGATTTCTTTATCTCCAGCAAC	1620								
QY	1686	TGTCCCATCAGTCTTCGAACTCAAATCTTACAGCAATGGGTACTTCTATGTCTTTATCAT	1745								
DB	1621	TGTCCCGTCCGCTTTTGAAGTCAAATCTTACAGCAATGGGTACTTCTATGTCTTTATCAT	1680								
QY	1746	GGAGGCCATCATAGCTTGCAGCTTTATGCAAGTCTGAAACAGAGGGACTGGGGGTCC	1805								
DB	1681	GGAGGCCATCATAGCTTGCAGCTTTATGCAAGTCTGAAACAGAGGGTTCGGAGGTCT	1740								
QY	1806	CATAGCACCCCACTTAACCTGATCAGCAGGAGCAGTGGTCCGGAAGCGGCCAGCCT	1865								
DB	1741	GCC---GGAGCCTGGCAACCTGATCAGCAGGAGCAGTGGTCCGGAAGCGGCCAGCCT	1797								
QY	1866	GTGCTACCTTCTCTCAATGAAGGCAACCATCTCACTGCCCTTGCAGACATTTTACCAAGT	1925								
DB	1798	GTGCTACCTTCTCTCAATGAAGGCAACCATCTCTGCCCTTGCAGACATTTTACCAAGT	1857								
QY	1926	CTGCCATGAAACAGTAGAGAAAGTTTATCCAGTATGGCAATTTTACAGTGGCAGAGCAGA	1985								
DB	1858	TTGTACAGAGACAGTAGAGAAAGTTTATCCAGTACCGAATTTCTACAGTGGCAGAGCAAG	1917								
QY	1986	TGACAGGAGAGATATCAGTCTCTAGTCTTCTGATCAGCAGCAGTGGGCAAGAGCTTCCAG	2045								
DB	1918	TGACAGGAGAGATATCAGTCTCTGCTGCTGCTGATCAGCAGCAGTGGGCAAGAGCTTCCAG	1977								
QY	2046	ACCTTTGTCTTGGAGAGTGTGAAGAGTGAAGAGTGAAGAGTGAAGTGGGGAGGAACAGCG	2105								
DB	1978	GCCTCTGAAGTGGAGAGTGAAGAGAGTGAAGAGTGAAGTGGGGAGGAACAGCG	2037								
QY	2106	AGATTGCTACCTGAAGGTGAGGCAATCCAAAGGAGCAACAGCAGTGTATCACCTTCTTACA	2165								

Db 2038 TGATTGCTACCTGAAGGTGAGCCAGGCCAAGGACACAGCAAGCAATTCATCACTTCTGTGCA 2097  
Qy 2166 GAGACTCCTTGGGCTTTGCTGGAGGCTCAGAGCTCTGCTGCCATCTTTGTTCACAACTT 2225  
Db 2098 GAGGCTTCTGGGGCCCTGCTAGAAGCTCAGAGCTCTGCTGCCATCTTTGTTCACACCTT 2157  
Qy 2226 CAGTGGTCTCTTCCAGAACCTCAGTATCTATCTGCAAAAGTTGCACAAATACCTAATAACGAG 2285  
Db 2158 CCGCGCGCCAGTCTCCGAGTCTGAGTACCTGCAGAAGCTGCACAGGTACCTTCTCACCCAG 2217  
Qy 2286 AACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCACATATTTGTTGGAAGAATGC 2345  
Db 2218 GACGGAGAGAACTGCTCGGGTGTACCTGAGAGTGCACATATCTGTTGGAAGAATGC 2277  
Qy 2346 TGTGAAATGTTTAAAGTATTTGGGGTTTCAAGGAGACCAAAAGAGAGAGTGTCTGT 2405  
Db 2278 TGTGAAATGTTTAAAGACATCGGGTTTCAAAGAGACCAAGCAGAGCGAGCTGT 2337  
Qy 2406 TTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACGACAAAACCTCTAGAAATATAT 2465  
Db 2338 CTTAGAACTGAGCAGCACTTTTCTACCTCAGCGGCGCGGAGAGCTCTCTGGAAATACAT 2397  
Qy 2466 TCTGAGTTTCTGTGTGCTG 2484  
Db 2398 TCTGAGCTTCGTGTGCTG 2416

RESULT 15  
AL391986/c  
LOCUS AL391986 196657 bp DNA linear PRI 30-NOV-2000  
DEFINITION Human DNA sequence from clone RP11-426E5 on chromosome 10, complete  
sequence.  
ACCESSION AL391986  
VERSION AL391986.12 GI:11544543  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 196657)  
Aehwell.R.

Direct Submission  
Submitted (30-NOV-2000) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
On Dec 4, 2000 this sequence version replaced gi:11414640.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence has been finished according to sequence map criteria  
as follows. An attempt is made to resolve all sequencing problems,  
such as compressions and repeats, but not necessarily within known  
annotated repeat sequence elements. Where the sequence is  
ambiguous, there is an annotation using the 'unsure' feature key.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 10, constructed by the Sanger Centre Chromosome 10  
Mapping Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr10  
RP11-426E5 is from the library RPCI-11.2 constructed at the Roswell  
Park Cancer Institute by the group of Pieter de Jong. For further  
details see http://bacpac.med.buffalo.edu/  
VECTOR: pBAC3.6  
IMPORTANT: This sequence is not the entire insert of clone  
RP11-426E5 it may be shorter because we sequence overlapping  
sections only once, except for a 100 base overlap.

The true left end of clone RP11-426E5 is at 3 in this sequence. The  
true right end of clone RP11-426E5 is at 196657 in this sequence.  
The true left end of clone RP11-32402 is at 133921 in this  
sequence. The true right end of clone RP11-159K22 is at 76098 in  
this sequence.

FEATURES  
source  
1..196657  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/clone="RP11-426E5"  
/clone\_id="RPCI-11.2"  
BASE COUNT 59606 a 42985 c 40722 g 53340 t  
ORIGIN  
Query Match 10.8%; Score 267.4; DB 9; Length 196657;  
Best Local Similarity 99.6%; Pred. No. 9.6e-62;  
Matches 268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1493 AGGGAATTGATCTCTCCACATTGGTCCGAAGACTTCTTTGTGATGAAAAGAGGAAGTCCCTGG 1552  
Db 50519 AGGGAATTGATCTCTCCACATTGGTCCGAAGACTTCTTTGTGATGAAAAGAGGAAGTCCCTGG 50460  
Qy 1553 CTGCTGATTTTGACCTCGGGGTTCTCAGGAAATTCAGAAGATGTAGTAATGCATGCCATAC 1612  
Db 50459 CTGCTGATTTTGACCTCGGGGTTCTCAGGAAATTCAGAAGATGTAGTAATGCATGCCATAC 50400  
Qy 1613 AGCTGCTGGGAAATTTGTCACATACCCACACTAGCAGGAACGATGAGTTTTTTATCA 1672  
Db 50399 AGCTGCTGGGAAATTTGTCACATACCCACACTAGCAGGAACGATGAGTTTTTTATCA 50340  
Qy 1673 CCCCAGCACAACTGTCCCATCAGTCTTCGAACCTCAACTCTTACAGCAATGGGGTACTTC 1732  
Db 50339 CCCCAGCACAACTGTCCCATCAGTCTTCGAACCTCAACTCTTACAGCAATGGGGTACTTC 50280  
Qy 1733 ATGCTTTTATCATGGAGGCCCATCATAGCT 1761  
Db 50279 ATGCTTTTATCATGGAGGCCCATCATAGGT 50251

Search completed: January 10, 2004, 04:05:57  
Job time : 5953.75 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 23:40:01 ; Search time 123.136 Seconds  
(without alignments)  
8903.931 Million cell updates/sec

Title: US-09-935-290-3  
Perfect score: 2484  
Sequence: 1 atgatgaatctgcactgac.....ttctgagttttgtggtgctg 2484

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	2.1	2598	US-09-252-991A-4323	Sequence 4323, Ap
C 2	53	2.1	2874	US-09-252-991A-4533	Sequence 4533, Ap
C 3	46.2	1.9	2589	US-09-328-352-2160	Sequence 2160, Ap
C 4	44.8	1.8	1830121	US-09-557-884-1	Sequence 1, Appli
C 5	44.8	1.8	1830121	US-09-643-990A-1	Sequence 1, Appli
C 6	38.6	1.6	1664976	US-08-916-421B-1	Sequence 1, Appli
C 7	37.2	1.5	1578	US-09-328-352-2557	Sequence 2557, Ap
C 8	36.6	1.5	2790	US-08-895-601-2	Sequence 2, Appli
C 9	35.6	1.4	7218	US-08-232-463-14	Sequence 14, Appl
C 10	35.2	1.4	1032	US-09-328-352-2440	Sequence 2440, Ap
C 11	34.8	1.4	202001	US-09-734-674-3	Sequence 3, Appli
C 12	34.2	1.4	825	US-09-328-352-3343	Sequence 3343, Ap
C 13	34	1.4	1711	US-08-749-903-2	Sequence 2, Appli
C 14	34	1.4	1711	US-09-088-641-2	Sequence 2, Appli
C 15	34	1.4	9046	US-08-227-536-1	Sequence 1, Appli
C 16	34	1.4	9046	PCT-US95-04682-1	Sequence 1, Appli
C 17	33.6	1.4	2774	US-08-466-548B-1	Sequence 1, Appli
C 18	33.6	1.4	2774	US-07-998-226F-1	Sequence 1, Appli
C 19	33.6	1.4	2774	PCT-US93-12560-1	Sequence 1, Appli
C 20	33.6	1.4	4821	US-08-913-374-1	Sequence 1, Appli
C 21	33.4	1.3	4403765	US-09-103-840A-2	Sequence 2, Appli
C 22	33.4	1.3	4411529	US-09-103-840A-1	Sequence 1, Appli
C 23	33.2	1.3	1534	US-08-858-207A-93	Sequence 93, Appl
C 24	33.2	1.3	1848	US-09-134-001C-447	Sequence 447, App
C 25	33.2	1.3	7672	US-09-220-132-24	Sequence 24, Appl
C 26	33	1.3	3752	US-08-961-527-208	Sequence 208, App
C 27	33	1.3	319608	US-09-539-333D-1	Sequence 1, Appli

28	33	1.3	319608	4	US-09-679-409-1	Sequence 1, Appli
C 29	32.8	1.3	2190	4	US-09-107-532A-3292	Sequence 3292, Ap
C 30	32.8	1.3	116592	4	US-09-818-512-3	Sequence 3, Appli
C 31	32.4	1.3	2327	4	US-09-149-476-107	Sequence 107, App
C 32	32.2	1.3	2290	4	US-09-620-312D-891	Sequence 891, App
C 33	32.2	1.3	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 34	32	1.3	2379	4	US-09-205-258-175	Sequence 175, App
C 35	32	1.3	10482	4	US-09-322-478-23	Sequence 23, Appl
C 36	32	1.3	20199	4	US-08-961-527-6	Sequence 6, Appli
C 37	32	1.3	148567	4	US-09-801-876B-3	Sequence 3, Appli
C 38	31.8	1.3	703	4	US-09-581-001B-3	Sequence 3, Appli
C 39	31.8	1.3	932	4	US-09-325-932A-28	Sequence 28, Appl
C 40	31.8	1.3	6080	4	US-09-620-312D-327	Sequence 327, App
C 41	31.8	1.3	7277	4	US-09-795-927-5	Sequence 5, Appli
C 42	31.8	1.3	1830121	4	US-09-557-884-1	Sequence 1, Appli
C 43	31.8	1.3	1830121	4	US-09-643-990A-1	Sequence 1, Appli
C 44	31.6	1.3	540	4	US-09-328-352-2250	Sequence 2250, Ap
C 45	31.6	1.3	6567	4	US-09-328-352-3637	Sequence 3637, Ap

ALIGNMENTS

RESULT 1  
US-09-252-991A-4323  
; Sequence 4323, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 4323  
; LENGTH: 2598  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4323

Query Match	2.1%;	Score 53;	DB 4;	Length 2598;
Best Local Similarity	59.7%;	Pred. No. 3.2e-06;		
Matches	89;	Conservative	0;	Mismatches 60; Indels 0; Gaps 0;
Qy	696	CCATATTGACTATCTGCTGCTCACTTTTCCTTCTGCGCATACATCAAGCACCATA	755	
Db	1029	CCACATGACTACTGCTGCTCTCTCTACTGCTGCTGCGCATCGCTCTGCGCGCGCGG	1088	
Qy	756	CATTGCTTCAGGCAATATCTCAACATCCCAATCTTCAGTACCTTGATCCATAAGCTTGG	815	
Db	1089	CATCGCGCGGGATCAACCTGAACATCGCGTCTGCTGCTGCGCATCGCTCTGCGCGCGCGG	1148	
Qy	816	GGGCTTCTTATACGACGAGGCTCGATG	844	
Db	1149	GGCCTTCTTATGCGCGCGAGCTTCAAGG	1177	

RESULT 2  
US-09-252-991A-4533/c  
; Sequence 4533, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18







NAME/KEY: misc feature  
LOCATION: (103998)..(103998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (148948)..(148948)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (163385)..(163385)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191989)..(191989)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191995)..(191995)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (231980)..(231980)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234220)..(234220)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature

LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1470091)..(1470091)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1664854)..(1664854)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-421B-1

Query Match 1.6%; Score 38.6; DB 4; Length 1664976;  
Best Local Similarity 50.8%; Pred. NO.14;

Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 338 TTTTATTCAAGAGCGAGATGTGCATAAGGGCATGTTGCCACCAATGTGACTGAAATG 397  
Db 541143 TCTTTATTAAACATAGAGAGGATAACAGATGCATGATATAGTAGTGAATGTCATG 541084  
QY 398 TCGTGAACAGCAGTAGAGTACAAGAGGCAATTCGAGAAGTGGCTGCTGAATTAACCCCTG 457  
Db 541083 CTCTTAAATGAAATATTAAAGAAATAACAAAATTTCTCAGTTGTAGGAGAGCCAA 541024  
QY 458 ATGGTTCTGCCAGCGCAATCAAAAGCGTTAAACAAAGTGAAGAAAGCTAAAGGA 517  
Db 541023 AAACCTCTGAAAGTTGGATACAATAATGGTGGTTAATGAAGAGCATGTGAAACTA 540964  
QY 518 T 518



```

; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-Fla
US-08-232-463-14

```

[illegible]

RESULT 10  
 US-09-328-352-2440/c  
 ; Sequence 2440, Application US/09328352  
 ; Patent No. 6562958  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 ; FILE REFERENCE: GTC99-03PA  
 ; CURRENT APPLICATION NUMBER: US/09/328,352  
 ; CURRENT FILING DATE: 1999-06-04  
 ; NUMBER OF SEQ ID NOS: 8252  
 ; SEQ ID NO 2440  
 ; LENGTH: 1032  
 ; TYPE: DNA  
 ; ORGANISM: Acinetobacter baumannii  
 US-09-328-352-2440

RESULT 11	
US-09-734	
; Sequenc	
; Patent	
; GENERAL	
; APPLIC	
; TITLE	
; TITLE	
; FILE R	
; CURREN	
; CURREN	
; NUMBER	
; SOFTWA	
; SEQ ID	
; LENGTH	
; TYPE:	
; ORGAN	
; FEATU	
; NAME/	
; LOCAT	
; OTHER	
US-09-734	

	Query Match 1.4%;	Score 34.8;	DB 4;	Length 202001;
	Best Local Similarity 60.6%;	Pred. No. 54;		
	Matches 57;	Conservative 0;	Mismatches 37;	Indels 0; Gaps 0;
Qy	2216	TTCACAACTTCAGTGGTCTCTTCAGAACCTGAGTATCTCGAAAGTTGCACAAATACC	2275	
Db	32955	TTCTCTAGTCTCTTAGGCCAGGTGTTAATGTTAGGATTTACAAAGTTGGTAAATAGA	33014	
Qy	2276	TAATATACCAAGACAGAAAGAAATGTTGCAGTATA	2309	
Db	33015	GAGAAACAGCAGAGAAATCGAAATGGCAGAGAAA	33048	

RESULT 12  
US-09-328-352-3343/c  
; Sequence 3343, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 3343  
; LENGTH: 825  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-3343

RESULT 13

US-08-749-903-2  
; Sequence 2, Application US/08749903  
; Patent No. 5759812  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hawkins, Phillip R.  
; TITLE OF INVENTION: NOVEL HUMAN SILENIUM-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/749,903  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PP-0163 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1711 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 5759812e  
; IMMEDIATE SOURCE:  
; CLONE: 989953  
; US-08-749-903-2

Query Match 1.4%; Score 34; DB 1; Length 1711;  
Best Local Similarity 63.4%; Pred. No. 3.4;  
Matches 52; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1258 GCTCTACTTTCCCTGGACGAGCGTTGTACACGCTATCTCTTCAAGACCCAGTGAT 1317  
Db 1270 GCTGTACAGTGCCTGGACAGCAGTTTACCCCTGATCTCATCAGGGAAGGCTCTGTGAT 1329

Qy 1318 GCTGCTGATGAGGTAGACACA 1339  
Db 1330 GCTGCAGGTTGATGTAGACACA 1351

RESULT 14  
US-09-088-641-2  
; Sequence 2, Application US/09088641  
; Patent No. 6312895  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hawkins, Phillip R.  
; TITLE OF INVENTION: NOVEL HUMAN SILENIUM-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/088,641  
FILING DATE:  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER: 08/749,903  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PP-0163 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6312895e  
IMMEDIATE SOURCE:  
CLONE: 989953  
US-09-088-641-2

Query Match 1.4%; Score 34; DB 4; Length 1711;  
Best Local Similarity 63.4%; Pred. No. 3.4;  
Matches 52; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1258 GCTCTACTTTCCCTGGACGAGCGTTGTACACGCTATCTCTTCAAGACCCAGTGAT 1317  
Db 1270 GCTGTACAGTGCCTGGACAGCAGTTTACCCCTGATCTCATCAGGGAAGGCTCTGTGAT 1329

Qy 1318 GCTGCTGATGAGGTAGACACA 1339  
Db 1330 GCTGCAGGTTGATGTAGACACA 1351

RESULT 15  
US-08-227-536-1/c  
; Sequence 1, Application US/08227536  
; Patent No. 5658784  
; GENERAL INFORMATION:  
; APPLICANT: Eckner, Richard  
; APPLICANT: Ewen, Mark  
; APPLICANT: Livingston, David  
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION  
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
; STREET: Ten Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/227,536  
; FILING DATE: 14-APR-1994  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Ph.D., Kathleen A.  
; REGISTRATION NUMBER: 34,380  
; REFERENCE/DOCKET NUMBER: DFCI-308XX

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-2290

TELEFAX: (617) 451-0313

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 9046 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cdna

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: CDS

LOCATION: 1200...8441

US-08-227-536-1

Query Match 1.4%; Score 34; DB 1; Length 9046;  
 Best Local Similarity 61.1%; Pred. No. 11;  
 Matches 55; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy	599	TCCTTTGGACATTCAAATTCACAAAGTCACACTTGAGATGGTTAAAGCTGCAACTGAGA	658
Db	4235	TGTTTCGGTAGATTCCATTTTACAGTCTTCCACTTTAGACTCTGAATATCTCCGGCTG	4176
Qy	659	CGAATTTGCCCGCTTCTGTTTCTACAGTTC	688
Db	4175	CGTATCTGCTGGTCTGGTTGATCCACTTC	4146

Search completed: January 10, 2004, 06:15:36  
 Job time : 138.136 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 22:59:07 ; Search time 3337.8 Seconds  
(without alignments)  
18087.435 Million cell updates/sec

Title: US-09-935-290-3

Perfect score: 2484

Sequence: 1 atgagtaactgcactgac.....ttctgagttttgtgtgtgctg 2484

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1843.6	74.2	3305	11 AK047295	AK047295 Mus muscu
2	848.2	34.1	918	13 B0437820	B0437820 AGENCOURT
3	760.2	30.6	1043	12 BQ062851	BQ062851 AGENCOURT
4	730.6	29.4	1055	12 BQ057957	BQ057957 AGENCOURT

5	703.8	28.3	965	10	BF794257
6	666.8	26.8	968	12	BQ058377
7	605	24.4	687	9	AL598680
8	564.2	22.7	637	14	CB268766
9	561.4	22.6	753	13	BU703927
10	547.2	22.0	580	10	AW976326
11	530	21.3	988	10	BG034520
12	511.4	20.6	776	14	CD000106
13	496.8	20.0	550	13	BQ694822
14	487.6	19.6	690	12	B1691948
15	467.8	18.8	825	13	BU263739
16	445	17.9	589	14	CA889049
17	430.2	17.3	799	10	BG704442
18	420	16.9	859	12	B1738244
19	418.2	16.8	746	14	BY741039
20	404.4	16.3	602	9	AL603562
21	398	16.0	446	12	BM414848
22	388.2	15.6	564	9	AA277375
23	381	15.3	1137	12	BM471397
24	369	14.9	553	9	AL596527
25	368.6	14.8	400	14	R73257
26	366.4	14.8	646	12	B1100769
27	365	14.7	573	14	CB269030
28	362.8	14.6	491	14	CA885724
29	359	14.5	535	14	CA896337
30	347.2	14.0	519	14	CA887606
31	347.2	14.0	527	14	CD554026
32	343.4	13.8	347	14	CB265571
33	340.2	13.7	571	9	AA172653
34	335.4	13.5	1037	10	BF037399
35	331.8	13.4	656	9	AA769055
36	329.8	13.3	411	12	BM146906
37	324	13.0	419	12	BM710854
38	321.8	13.0	823	14	CD362496
39	320.4	12.9	1310	14	CD502509
40	314.8	12.7	889	10	BF783411
41	314.6	12.7	751	14	CA384438
42	311.2	12.5	465	9	AA681924
43	306.6	12.3	713	13	BU444448
44	302.2	12.2	400	12	BG993243
45	300.8	12.1	635	12	BU063511

#### ALIGNMENTS

#### RESULT 1

AK047295

LOCUS

DEFINITION

AK047295 3305 bp mRNA linear HTC 05-DEC-2002  
Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length  
enriched library, clone:B930046K04 product:glycerol-3-phosphate  
acyltransferase, mitochondrial, full insert sequence.

ACCESSION

AK047295

VERSION

AK047295.1 GI:26092071

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1

AUTHORS

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

JOURNAL

MEDLINE

PUBMED

REFERENCE

2

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

PUBMED

11042159



REFERENCE AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL MEDLINE PUBMED	Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861
REFERENCE AUTHORS	4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaïdo, I., Pesole, G., Saito, M., Quackenbush, J., Schriml, L.M., Staab, J., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bernaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL MEDLINE PUBMED	Nature 409 (6821), 685-690 (2001) 21085660 11217851
REFERENCE AUTHORS	5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL MEDLINE PUBMED	Nature 420, 563-573 (2002) 6 (bases 1 to 3305)
REFERENCE AUTHORS	Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayaehida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, K., Saiton, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
FEATURES	Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers

721 TTCAATTCCTCTTGTGCATAAATCAATCAAAAGCACCATATACATTTGCTTCAGGCAATAATCTCAAC 780  
Db |||||  
876 TTCAATCTCTTTTGGCACAACATCAAGGCGCGTACATCGCTCGGGCAATAATCTCAAC 935  
Qy |||||  
781 ATCCCAATCTTCAGTACCTTGTATCATTAAGCTTTGGGGCTTCTTCATACGACGAGGCTC 840  
Db |||||  
936 ATCCCGGTCTTCAGTACCTTGTATCATTAAGCTTTGGGGCTTCTTCATTAAGACGAGGCTC 995  
Qy |||||  
841 GATGAAACACCATGATGAGCGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db |||||  
996 GATGAAACCCAGATGAGCGCAACATCTGTACAGAGGTTGCTCATGCGCATGTA 1055  
Qy |||||  
901 GTTGAATTTCTCGACAGAGCAATCTTGTGAGATCTTCTGGAAGGACACGTTCTAGG 960  
Db |||||  
1056 GTTGAATCTCTCCGACAGCAGCAGTCTCTGTGAGATCTTCTGGAAGGACACCGCTCCGC 1115  
Qy |||||  
961 AGTGAAACCTCTGTGCTCGGCAGGACTTTGTCTAGTGTGGTAGACTCTGTCT 1020  
Db |||||  
1116 AGTGCAAGACCTCTGTGCGCGGCGAGGCTCTCTCAGTGGTAGTGGTACTCTGTGC 1175  
Qy |||||  
1021 ACCAATGTCATCCAGACATCTTGATAATACCTGTGTGAATCTCCTATGATCGCATATC 1080  
Db |||||  
1176 TCCAACACCATCCCGCAATCTCTGTATACCCGTGGGCATCTCGTATGATCGCATATC 1235  
Qy |||||  
1081 GAAGTCTACTAATGTTGTAACAACTGGGCAACCTAAGAAAGATGAGAGCTCTGGAGT 1140  
Db |||||  
1236 GAAGTCTACTAATGTTGTAACAACTGGGCAACCTAAGAAAGATGAGAGCTCTGGAGT 1295  
Qy |||||  
1141 GTAGCAAGAGTGTATTAGATGTTACGAAACCTATGTTGTTGTCGAGTGGATTTT 1200  
Db |||||  
1296 GTGGCAGAGGCTTATCAGAAATGTCGCGGAAACCTACGGCTACGTCGAGTGGATTTT 1355  
Qy |||||  
1201 GCACAGCATTCTCTTAAAGAAATATTTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCT 1260  
Db |||||  
1356 GCACAGCATTCTCTTGAAGAAATATTTAGAAAGCCAAAGTCAGAAACCTGTATCTGCC 1415  
Qy |||||  
1261 CTACTTTCCCTGGAGCAAGCGTTGTACGCTATATCTTCTTCAAGACCCAGTGTGCT 1320  
Db |||||  
1416 CCCCTTTCTCTGGAGCAAGCACTGTTTACCAGCATCTCTTCTTCAAGACCCGAATGATT 1475  
Qy |||||  
1321 GCTGATGAGGTAGAGACAGCTCCATTAATGATCCAGAAATGCAACAGATGAATCCCTA 1380  
Db |||||  
1476 GCTGATGAACATCAAGACTTATCCAGTAACAGTTCAGAAACCCAGCAGACGAAAGCTTC 1535  
Qy |||||  
1381 CGAAGGAGGTTGATTGCAAAATCTGCTCAGCATATCTTATTCATCTGCTAGCAAGTCTGT 1440  
Db |||||  
1536 CGAAGGCTGATTGCAAACTGGCTGAGCAGATTTCTTTCACCGCCAGCAAGTCTGTC 1595  
Qy |||||  
1441 GCCATTATGTCACACACATTTGCTGCTTGCCTCTCTACAGACACAGCAGGGAAT 1500  
Db |||||  
1596 GCTATCATGTCACCCCAATTTGTGGCTGTCTGCTCTCTACAGACACAGCAGGGAATC 1655  
Qy |||||  
1501 GATCTCTCCATTTGGTCCAGAGCTTCTTTGTGATGAAAGAGGAGTCTCGGCTCGTGAAT 1560  
Db |||||  
1656 CATCTCTCCAGCTTGTGGAAGCTTCTTTGTGATGAAGAGGAGGAGTCTTACGCTCGCAT 1715  
Qy |||||  
1561 TTTGACCTGGGTTCTCAGGAAATTCAGAGATGTAGTAATGATGCTATGCCATACAGTCTGT 1620  
Db |||||  
1716 TTGACCTAGGCTTCTCGGGAATTCAGAAATGTGTCATGTCATGCTATTCAGCTCTGT 1775  
Qy |||||  
1621 GGAATTTGTGTACAAATCACCCACACTAGCAGGAAACGATGATGTTTTTATCACCCCGAGC 1680  
Db |||||  
1776 GGAATTTGTGTACAAATCACCCACACTAGCAGGAAACGATGATGTTTTTATTCACCCGAGC 1835  
Qy |||||  
1681 ACACTGTCCCATCAGTCTTCCAACTCAACTTCTACAGCAATGGGTTACTTATGCTCTTT 1740  
Db |||||  
1836 ACACTGTCCCGTCACTTTTGAATCACTCACTTCTACAGCAATGGGTTACTTATGTTGTC 1895  
Qy |||||  
1741 ATCATGGAGGCATCATAGCTTGCAGCTTTTATGAGTCTTCTGAAACAGAGGGGACTGGG 1800  
Db |||||  
1896 ATCATGGAGGCATCATAGCTTGCAGCTTTTATGAGTCTTCTGAAATCAAGAGGTTCTCTGA 1955

Qy 1801 GGTCCCACTAGCACCCCACTTAACCTGATCAGCAGGAGCAGCTGGTGCAGAAAGCGCGCC 1860  
Db |||||  
1956 GGGTCGCTGAGGCGCTCGGCAACCTGATCAGCAGGAGCAGCTGGTGCAGAAAGCGCGCC 2015  
Qy |||||  
1861 AGCCTGTCTACCTTCTCTCCAAATGAAGGACCATCTCAGTGCCTTGGCCACACATTTTAC 1920  
Db |||||  
2016 AGCCTGTCTACCTTCTCTTAACGAAGGTACCAATTTCTCTGCGCTGCGCAGATTTTAC 2075  
Qy |||||  
1921 CAACTCTGCCATGAAAACAGTAGGAAAGTTTATCCAGTATGGCATTTCTTACAGTGGCAGAG 1980  
Db |||||  
2076 CAACTTTGTCTATGAGACAGTTGGCAAGTTCTATCCAGTATGGCATTTCTCAGTGGCAGA- 2134  
Qy |||||  
1981 CACGATCACCCAGGAAGATATCAGTCTCTAGTCTTCTCAGCAGCAGTGGGCAAGAAAGCTT 2040  
Db |||||  
2135 ---GATCACCCAGGAAGATGTCACTCTGGCTTGCAGAACAGCAGTGGGCAAGAAAGCTT 2191  
Qy |||||  
2041 CCAGAACTTTGTCTTGGAGAAAGTGAAGAAGATGAAGACAGTACTTTTGGGGAGAA 2100  
Db |||||  
2192 CCGGAAC---TGAACTGGAGAAAGTGCAGGAAAGATGAAGACAGTACTTTTGGTGGAGAG 2248  
Qy |||||  
2101 CAGCGAGATTGCTACCTGAAGGTGAGCCAAATCCAAAGGAGCACAGCAGTATTTATCACCTTC 2160  
Db |||||  
2249 CAGCGAGATTGCTATCTCAAGGTGAGCAGTCCAAAGGAGCACAGCAATTTATCACCTTT 2308  
Qy |||||  
2161 TTACAGAGACTCTTGGGCCCTTTGCTGAGGCCCTACAGCTCTGCTGCCATCTTTGTTTAC 2220  
Db |||||  
2309 CTTCAGAGGCTTCTAGTCCCTCTAGAAAGCTTACAGCTCTGCTGCCATCTTTGTTCCAC 2368  
Qy |||||  
2221 AACTTCACTGCTCTGTTTCAGAAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAATA 2280  
Db |||||  
2369 AACTTCACTGCTCTGTTTCAGAAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAATA 2428  
Qy |||||  
2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTCAGAGTGCCACATATGTTCTGTGAAG 2340  
Db |||||  
2429 ACCAGAACAGAAAGAAAGTTCGCGTGTAGCTGAGAGTGCCACATATGTTCTGTGAAG 2488  
Qy |||||  
2341 AATGCTGTGAAAATGTTTAAAGGATATTTGAGGATATTTTCAAGGAGACAAACAAAGAGAGTG 2400  
Db |||||  
2489 AACGCTGTGAAAATGTTTAAAGGATATTTTAAAGGATATTTTCAAGGAGACAAACAAAGAGTG 2548  
Qy |||||  
2401 TCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAATGCAACCGCAACAAACTTCTAGAA 2460  
Db |||||  
2549 TCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAGTGTCAACCGGCAAACTCTCTAGAG 2608  
Qy |||||  
2461 TATATCTGAGTTTGTGGTGTCTG 2484  
Db |||||  
2609 TATATCTGAGTTTGTGGTGTCTG 2632

## RESULT 2

BQ437820

LOCUS

DEFINITION BQ437820 918 bp mRNA linear EST 24-MAY-2002  
AGENCOURT 7917014 NIH\_MGC\_72 Homo sapiens cdna clone IMAGE:6161400  
5', mRNA sequence.

ACCESSION

BQ437820

VERSION

BQ437820.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 918)

AUTHORS

NIH-MGC http://mgi.nci.nih.gov/

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgabs-x@mail.nih.gov  
Tissue Procurement: ATCC/DCTP/DMP  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov									
Plate: L1AM13513 row: k column: 01									
High quality sequence stop: 657.									
FEATURES									
source									
1. .918									
/organism="Homo sapiens"									
/mol_type="mRNA"									
/db_xref="taxon:9606"									
/clone="IMAGE:6161400"									
/tissue_type="melanotic melanoma"									
/lab_host="DH10B (phage-resistant)"									
/clone_lib="NIH MGC_72"									
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 2 kb. Library constructed by Life Technologies."									
BASE COUNT 244 a 223 c 221 g 224 t 6 others									
ORIGIN									
Query Match 34.1%; Score 848.2; DB 13; Length 918;									
Best Local Similarity 98.7%; Pred. No. 1.1e-227;									
Matches 906; Conservative 0; Mismatches 6; Indels 6; Gaps 5;									
Qy	1291	GCTATATCTCTTCAAGACCCAGTGATGCTGCTGATGAAGGTAGACACGTCCTAAT	1350						
Db	1	GCTATATCTCTTCAAGACCCAGTGATGCTGCTGATGAAGGTAGACACGTCCTAAT	60						
Qy	1351	GAGTCCAGAATGCAACAGATGAATCCTCAGGAAGAGTTGATTGCAAAATCTGGCTGAG	1410						
Db	61	GAGTCCAGAATGCAACAGATGAATCCTCAGGAAGAGTTGATTGCAAAATCTGGCTGAG	120						
Qy	1411	CATATTCTATTCACTGCTAGCAAGTCTGTGCCATTATGTCCACACACATTTGGCTTGC	1470						
Db	121	CATATTCTATTCACTGCTAGCAAGTCTGTGCCATTATGTCCACACACATTTGGCTTGC	180						
Qy	1471	CTGCTCCTCTACAGACACAGCAGGGAATTGATCTCTCCACATTTGGTTCGAAGACTTCTTT	1530						
Db	181	CTGCTCCTCTACAGACACAGCAGGGAATTGATCTCTCCACATTTGGTTCGAAGACTTCTTT	240						
Qy	1531	GTGATGAAGAGGAAGTCTGGCTCGTGATTTTTGACCTGGGGTCTCAGGAAATTCAGAA	1590						
Db	241	GTGATGAAGAGGAAGTCTGGCTCGTGATTTTTGACCTGGGGTCTCAGGAAATTCAGAA	300						
Qy	1591	GATGTAGTAATGATGTCATACAGCTGCTGGGAAATTTGTGTCAAAATCACCACACTAGC	1650						
Db	301	GATGTAGTAATGATGTCATACAGCTGCTGGGAAATTTGTGTCAAAATCACCACACTAGC	360						
Qy	1651	AGGAACGATGAGTTTTTTATCACCCCCAGCACAACTGTCCCAATCAGTCTTCGAACTCAAC	1710						
Db	361	AGGAACGATGAGTTTTTTATCACCCCCAGCACAACTGTCCCAATCAGTCTTCGAACTCAAC	420						
Qy	1711	TTCTACAGCAATGGGGTACTTTCATGCTTTATCATGGAGGCCATCATAGCTTGCAGCCTT	1770						
Db	421	TTCTACAGCAATGGGGTACTTTCATGCTTTATCATGGAGGCCATCATAGCTTGCAGCCTT	480						
Qy	1771	TATGCAGTCTCTGAACAAGAGGGGACTGGGGGTCCCACTAGCACCCACACTAACTGATC	1830						
Db	481	TATGCAGTCTCTGAACAAGAGGGGACTGGGGGTCCCACTAGCACCCACACTAACTGATC	540						
Qy	1831	AGCCAGGAGCAGCTGGTTCGGAAGGGCCAGCCTGTGCTACCTTCTCTCCAATGAAGGC	1890						
Db	541	AGCCAGGAGCAGCTGGTTCGGAAGGGCCAGCCTGTGCTACCTTCTCTCCAATGAAGGC	600						
Qy	1891	ACCATCTCACTGCTTCGCAGACATTTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTT	1950						
Db	601	ACCATCTCACTGCTTCGCAGACATTTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTT	660						
Qy	1951	ATCCAGTATGGCATTTCTTACAGTGGCAGACCAAGTACACAGG-AAAGATATCACTCTCAG	2009						
Db	661	ATCCAGTATGGCATTTCTTACAGTGGCAGACCAAGTACACAGGAAAGATATCACTCTCAG	720						
Qy	2010	TCTTGTCTGAGCAGCAGTGGGCAAGAAGCTTTCAGAACTTTT-GTCTTGGAGAGTATGATG	2068						

Db 232 GAATACAGTGTGGTGGATGTAAGCACACACAGTGAGGAATGGGTGAGTGCGCTTTTGA 291

Qy 121 CCCACCGTCTTCAGATCTGCACTTTAAATGGAAGAAAGCCTAATCAGTCGGAAGG 180

Db 292 CCCACCGTCTTCAGATCTGCACTTTAAATGGAAGAAAGCCTAATCAGTCGGAAGG 351

Qy 181 CCATTGTTGGAAGATGTTTACTCTGCACTCCCAAGAGTGGGACAAATTTTCAAC 240

Db 352 CCATTGTTGGAAGATGTTTACTCTGCACTCCCAAGAGTGGGACAAATTTTCAAC 411

Qy 241 CCCAGTATCCCGTCTTTGGGTTGGGAATGTTATTTATATCAATGAACCTCACACAGA 300

Db 412 CCCAGTATCCCGTCTTTGGGTTGGGAATGTTATTTATATCAATGAACCTCACACAGA 471

Qy 301 CACCGCGATGGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTG 360

Db 472 CACCGCGATGGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTG 531

Qy 361 CATAGGCGATGTTTGCACCAATGTGCTGAATAATGTCTGAACAGCAGTAGAGTACAA 420

Db 532 CATAGGCGATGTTTGCACCAATGTGCTGAATAATGTCTGAACAGCAGTAGAGTACAA 591

Qy 421 GAGGCAATGTCAGAGTGGCTGCTGAATTAACCTGATGTTCTGCCCAGCAGCAATCA 480

Db 592 GAGGCAATGTCAGAGTGGCTGCTGAATTAACCTGATGTTCTGCCCAGCAGCAATCA 651

Qy 481 AAAGCCGTTAAACAAAGTGAAGAAAGAGCTTAAAGGATTTTCAAGAAATGGTTGCCACT 540

Db 652 AAAGCCGTTAAACAAAGTGAAGAAAGAGCTTAAAGGATTTTCAAGATATGGTTGCCACT 711

Qy 541 GTCTCACCGCAATGATCAGATGATGCTGGGTGGGTGCTGCTGCTTAAACTGTTCAACAGCTTC 600

Db 712 GTCTCACCGCAATGATCAGATGATGCTGGGTGGGTGCTGCTGCTTAAACTGTTCAACAGCTTC 771

Qy 601 TTTTGGACATTCAAATTCACAAAGTCAAAAGTCAA-CTTGAGATGTTAAAGCTGCACTGAGAC 659

Db 772 TTTTGGACATTCAAATTCACAAAGTCAA-CTTGAGATGTTAAAGCTGCACTGAGAC 831

Qy 660 GAAATGCGCGTCTCTGTTTCTACCAAGTTCATAGATCCCATATGACTATCT-GCTGCTCA 718

Db 832 GAAATGCGCGTCTCTGTTTCTACCAAGTTCATAGATCCCATATGACTATCTGCTGCTCA 891

Qy 719 CTTCATTTCTCTTGGCATAACATCAAGACCAATACATGCTTTCAGGCAATATCTCA 778

Db 892 CTTCATTTCTCTTGGCATAACATCAAGACCAATACATGCTTTCAGGCAATATCTCA 951

Qy 779 ACATCCCAATCTTCAGTACCTTGATCCATAGCTT--GGGGCTTCTTCATACAGCAGAG 836

Db 952 ACATCCCAATCTTCAGTACCTTGATCCATAGCTT--GGGGCTTCTTCATACAGCAGAG 1011

Qy 837 GCTCGA-TGAACACACAGATGACGGAAGA 866

Db 1012 GCTCCATTTGAACCCCAATGACGGTAA 1042

RESULT 4  
LOCUS BQ057957  
DEFINITION BQ057957 1055 bp mRNA linear EST 29-MAR-2002  
5', mRNA sequence.  
ACCESSION BQ057957  
VERSION BQ057957.1 GI:19817297  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1055)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2066 row: d column: 21  
High quality sequence stop: 666.

#### FEATURES

source  
Location/Qualifiers  
1..1055  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:5813732"  
/tissue\_types="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT 285 a 248 c 240 g 281 t 1 others  
ORIGIN

Query Match 29.4%; Score 730.6; DB 12; Length 1055;  
Best Local Similarity 93.5%; Pred. No. 1.8e-194;  
Matches 817; Conservative 0; Mismatches 50; Indels 7; Gaps 5;  
Qy 1 ATGATGAATCTGACATGACCCCTTGGTACATAGATGTTCTTATCTGCCACATTCATCA 60  
Db 170 ATGATGAATCTGACATGACCCCTTGGTACATAGATGTTCTTATCTGCCACATTCATCA 229

Qy 61 GAATACAGTGTGGTTCGATGTAAGCACACAGTCAGGAATGGGTGAGTGCGCTTTAGA 120  
Db 230 GAATACAGTGTGGTTCGATGTAAGCACACAGTCAGGAATGGGTGAGTGCGCTTTAGA 289

Qy 121 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCTTAATCAGTCGGAAGG 180  
Db 290 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCTTAATCAGTCGGAAGG 349

Qy 181 CCATTGTTGGAAGATGTTGTTACTCTGCACTCCCAAGAGTGGGACAAATTTTCAAC 240  
Db 350 CCATTGTTGGAAGATGTTGTTACTCTGCACTCCCAAGAGTGGGACAAATTTTCAAC 409

Qy 241 CCCAGTATCCCGTCTTTGGGTTGGGAATGTTATTTATATCAATGAACCTCACACAGA 300  
Db 410 CCCAGTATCCCGTCTTTGGGTTGGGAATGTTATTTATATCAATGAACCTCACACAGA 469

Qy 301 CACCGCGATGGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTG 360  
Db 470 CACCGCGATGGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTG 529

Qy 361 CATAGGCGATGTTTGCACCAATGTGCTGAATAATGTCTGAACAGCAGTAGAGTACAA 420  
Db 530 CATAGGCGATGTTTGCACCAATGTGCTGAATAATGTCTGAACAGCAGTAGAGTACAA 589

Qy 421 GAGGCAATGTCAGAGTGGCTGCTGAATTAACCTGATGTTCTTCCCGCAGCAGCAATCA 480  
Db 590 GAGGCAATGTCAGAGTGGCTGCTGAATTAACCTGATGTTCTTCCCGCAGCAGCAATCA 649

Qy 481 AAAGCCGTTAAACAAAGTGAAGAAAGAGCTTAAAGGATTTCTTCAAGAAATGGTTGCCACT 540  
Db 650 AAAGCCGTTAAACAAAGTGAAGAAAGAGCTTAAAGGATTTCTTCAAGAAATGGTTGCCACT 709

Qy 541 GTCTCACCGCAATGATCAGATGATGCTGGGTGGGTGCTGCTTAAACTGTTCAACAGCTTC 600  
Db 710 GTCTCACCGCAATGATCAGATGATGCTGGGTGGGTGCTGCTTAAACTGTTCAACAGCTTC 769

```
QY 601 TTTTGAACATTCAAAATTCACAAAGGTCAACTTTGAGATGGTTAAAGCTGCAACTGAGACG 660
|||||
Db 770 TTTTGGACATTCAA-TCACAAAGGTCAACCTGAGATGGTTAAAGCTGCAACTGAGAG 828

QY 661 AATTGCGCGTCTGTTTACAGTTTCATAGATCCATATGACATATCTGCT-GCTAC 719
|||||
Db 829 AATTGCGCGTCTGTTTACAGTTTCATAGATCCATATGACATATCTGCTGCTGCTAC 888

QY 720 TTTTCAATCTCTTCTGCGCATTAACATCAAGACCATATGCTTTCAGGCAATATCTCAA 779
|||||
Db 889 TTTTCTCTCTCTGCGCATTAACATCAAGGCGCTTTCATTGCTTTCAGGCAATATCTCAA 948

QY 780 CATCCCAAT-CTTCAGTACCTTGAT--CCATAAGCTTGGGGGCTTC--TTCAATACGACGA 834
|||||
Db 949 CATCCCAATCTTCAGTACCTTGATTTGCTTCCCTAAACCTTGGGGGCTTCTCTTACACAA 1008

QY 835 AGGCTCGATGAACACACAGATGACGGAAGATG 868
|||||
Db 1009 AGGCTCGTGAACACCCCAAGTGGGACGGAAG 1042

RESULT 5
BF794257
LOCUS 60225455F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4338997 5',
DEFINITION mRNA sequence.
ACCESSION BF794257
VERSION BF794257.1 GI:12099311
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM948 row: m column: 14
High quality sequence stop: 696.
FEATURES
    source
        1..965
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4338997"
            /tissue_type="lymphoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_85"
            /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
            Average insert size 1.867 kb. Library enriched for
            full-length clones and constructed by Life Technologies.
            Note: this is a NIH_MGC Library."
BASE COUNT 273 a 222 c 226 g 244 t
ORIGIN
Query Match 28.3%; Score 703.8; DB 10; Length 965;
Best Local Similarity 97.6%; Pseq No. 6 6e-187;
Matches 758; Conservative 0; Mismatches 12; Indels 7; Gaps 4;

QY 964 GGAAAAACCTCTTGTGCTCGGGCAGGACTTTTGTCACTGTGGTAGATCTGTCTACC 1023
|||||
Db 1 GGAAAAACCTCTTGTGCTCGGGCAGGACTTTTGTCACTGTGGTAGATCTGTCTACC 60
```

RESULT 6  
BQ058377  
LOCUS  
DEFINITION  
5', mRNA sequence.

BQ058377  
AGENCOURT 6794559 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5814217  
5', mRNA sequence.

ACCESSION  
VERSION  
BQ058377  
BQ058377.1 GI:19817717  
EST.

KEYWORDS  
SOURCE  
Homo sapiens (human)  
Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM2067 row: i column: 02

High quality sequence stop: 722.

# FEATURES

Location/Qualifiers

1..968

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5814217"

/tissue type="lymphoma, cell line"

/lab host="DH10B (phage-resistant)"

/clone lib="NIH\_MGC\_99"

/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dr priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGCAG(G). Size-selected >500bp for average insert size

1.8kb. Library constructed by Ling Hong in the laboratory

of Gerald M. Rubin (University of California, Berkeley)

using ZAP-cDNA synthesis kit (Stratagene) and Superscript

II RT (Life Technologies). Note: this is a NIH\_MGC

Library."

276 a 214 c 216 g 262 t

## BASE COUNT

ORIGIN

Query Match 26.8%; Score 666.8; DB 12; Length 968;

Best Local Similarity 96.8%; Pred. No. 1.9e-176;

Matches 724; Conservative 0; Mismatches 17; Indels 7; Gaps 4;

Qy 1 ATGATGAATCTGCATGACCTGGTACATAGATGTTCTTATCTGCCACATTCATCA 60

Db 170 ATGATGAATCTGCATGACCTGGTACATAGATGTTCTTATCTGCCACATTCATCA 229

Qy 61 GAATACAGTGTGGTGCAGTGAAGCACACAAAGTGAAGTGGGTGAGTGGCTTTAGA 120

Db 230 GAATACAGTGTGGTGCAGTGAAGCACACAAAGTGAAGTGGGTGAGTGGCTTTAGA 289

Qy 121 CCCACCGTCTTCAGATCTGCACCTTTAAATGGAAAGAAAGCCTAATGAGTCGGAAGAAG 180

Db 290 CCCACCGTCTTCAGATCTGCACCTTTAAATGGAAAGAAAGCCTAATGAGTCGGAAGAAG 349

Qy 181 CCATTTGTGAAGATGTTGTTACTCTGCTGCTCCAGAGCTGGGACAAATTTTCAAC 240

Db 350 CCATTTGTGAAGATGTTGTTACTCTGCTGCTCCAGAGCTGGGACAAATTTTCAAC 409

Qy 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTATTTATATCAATGAACCTCACACAAGA 300

Db 410 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTATTTATATCAATGAACCTCACACAAGA 469

Qy 301 CACCGGATGCGTTGCAAGACGCTTTCTTACGTTCTTTTATTCAGAGCGAGATGTG 360

Db 470 CACCGGATGCGTTGCAAGACGCTTTCTTACGTTCTTTTATTCAGAGCGAGATGTG 529

Qy 361 CATAGGCGATGTTGGCCACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

Db 530 CATAGGCGATGTTGGCCACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 589

Qy 421 GAGGCAATTCAGAGTGGCTGCTGAATTAACCTGATGTTCTGCCAGCAGCAATCA 480

Db 590 GAGGCAATTCAGAGTGGCTGCTGAATTAACCTGATGTTCTGCCAGCAGCAATCA 649

Qy 481 AAAGCGTTAAAGAGTGAAGAAAGAGCTTAAAGAGTCTTCAAGAAATGGTTGCCACT 540

Db 650 AAAGCGTTAAAGAGTGAAGAAAGAGCTTAAAGAGTCTTCAAGAAATGGTTGCCACT 709

Qy 541 GTCTCAGCGGAATGATCAGACTGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 600

Db 710 GTCTCAGCGGAATGATCAGACTGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 769

Qy 601 TTTTGGAACTTCAAAATTCACAAAGGTCACTTCAGATGGTTAAAGCTGCAACT---GAG 657

Db 770 TTTTGGAACTTCAAAATTCACAAAGGTCACTTCAGATGGTTAAAGCTGCAACTGGAGAC 829

Qy 658 ACGAATTTCGCGCTTCTGTTTCTACCAATT--CATAGATCCCATATTTGACTAT-CTGCTG 714

Db 830 GAAATTTGCCCGCTTGGTTTCTACCAATTTCATAGATCCCATATTTGACTATCTCTGCTG 889

Qy 715 CTCAC-TTTCATTTCTTCTGCCATAAC 741

Db 890 CTCAC-TTTCATTTCTTCTGCCATAAC 917

RESULT 7

LOCUS

DEFINITION

AL598680

DKFZp313P1521\_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone

AL598680

AL598680.1 GI:15161371

EST.

ORGANISM

Homo sapiens (human)

REFERENCE

1 (bases 1 to 687)

Poustka,A., Wellenreuther,R., Mewes,H.W., Weil,B. and Wiemann,S.).

EST (Poustka,A., Wellenreuther,R., Mewes,H.W., Weil,B. and Wiemann

,S.).

Unpublished

Contact: Poustka A.J.

Department Lehrach

Max-Planck-Institute for Molecular Genetics

Innestrasse 73, 14195 Berlin, Germany

Tel: +49-30-84131623

Fax: +49-30-84131128

Email: poustka@mpg-berlin-dahlem.mpg.de

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.

No sl sequence available.

This clone (DKFZp313P1521) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Location/Qualifiers

1..687

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="DKFZp313P1521"

/dev stage="adult"

/lab\_host="DH10B"

/clone\_lib="313 (synonym: hlcc2)"

/note="vector: pTriplex2; Site 1: SfiIA; Site 2: SfiIB;

cDNA-collection"

BASE COUNT 202 a 139 c 168 g 178 t

ORIGIN

Query Match 24.4%; Score 605; DB 9; Length 687;

Best Local Similarity 100.0%; Pred. No. 4.5e-159;

Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1880 CCAATGAAGCACCCTCTCACTGCTTGCAGACATTTTACCAGTTCGCCATGAACAG 1939

Db 1 CCAATGAAGCACCCTCTCACTGCTTGCAGACATTTTACCAGTTCGCCATGAACAG 60

Qy 1940 TAGAAGATTTATCCAGTATGGCATTTTACAGTGGCAGAGCAGATGACCAAGATA 1999

Db 61 TAGAAGATTTATCCAGTATGGCATTTTACAGTGGCAGAGCAGATGACCAAGATA 120

Qy 2000 TCAGTCTCTAGTCTTGTCAGCAGCAGTGGGACAAAGAGCTTCAGAACCTTTGCTTGA 2059

Db 121 TCAGTCTCTAGTCTTGTCAGCAGCAGTGGGACAAAGAGCTTCAGAACCTTTGCTTGA 180







This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

Location/Qualifiers

source

1. 753  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CS7BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE: 6406167"  
/tissue type="whole brain"  
/dev stage="embryo 12.5dpc"  
/lab host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH BMAP\_F00"  
/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TCAGAGAGCC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 196 a 202 c 173 g 180 t 2 others

ORIGIN

Query Match 22.6%; Score 561.4; DB 13; Length 753;

Best Local Similarity 84.9%; Pred. No. 9.7e-147;

Matches 639; Conservative 0; Mismatches 113; Indels 1; Gaps 1;

QY 1033 CCAGACATCTTGAATAACCTGTTGGATCTCTTATGATCGCATATCGAAGGTCACCTAC 1092  
DB 1 CCCGACATCCCTCGTCATCCCGTGGGCATCTCGTATGATCGCATATCGAAGGTCACCTAC 60  
QY 1093 AATGTGTAACAA-CTGGGCAACCTAAGCAAGATGAGACCTGTGGAGTGTAGCAAGG 1151  
DB 61 AATGCGCAACAGTTGGGAAGCCCAAGAAACAGAGAGCTCTGGAGTGTGGCGAGAG 120  
QY 1152 TGTATTAGAAATGTTACAAAAAACTATGGTTGTCTCCAGTGGATTTGCACAGCCATT 1211  
DB 121 CGTTATCAGATGTCGCGAAATACTACGGCTACGTCGAGTGGATTTGCACAGCCATT 180  
QY 1212 TTCCTTAAGGAATATTTAGAAAGCAAGTACAGAAACCGGTGTCTGCTCTACTTTCCCT 1271  
DB 181 TTCCTTGAAGGAATATTTAGAAAGCCAGATCAGAAACCTGTATCTGCCCCCTTTCTCT 240  
QY 1272 GGAGCAAGCTGTTTACCAGCTATCTTCTTCCAGCCAGTATGTCGTGATGAGG 1331  
DB 241 GGAGCAAGCTGTTTACCAGCTATCTTCTTCCAGCCAGTATGTCGTGATGAGG 300  
QY 1332 TAGACACAGCTCCATTATAGTCCAGAAATGCAACAGATCAATCCCTTACCAAGAGGTT 1391  
DB 301 TCAAGACCTATCCAGTACAGATCCAGAAACCCAGCAGACAGAGCTTCCAGCAGAGCT 360  
QY 1392 GATTGCAAACTGGCTGAGCATATTTATTTCACTGTCTAGCAAGTCTGTGCCATTTATGTC 1451  
DB 361 GATTGCAAACTGGCTGAGCATATTTATTTCACTGTCTAGCAAGTCTGTGCCATTTATGTC 420  
QY 1452 CACACACATTTGGCTTGCCTCTCTACAGACAGAGGAGGAATGATCTCTCCAC 1511  
DB 421 CACACACATTTGGCTTGCCTCTCTACAGACAGAGGAGGAATGATCTCTCTCCAC 480  
QY 1512 ATTGGTCAAGACTTCTTTGTGTAAGAGAGAGTCTCTGCTCGTGTATTTGACCTGGG 1571  
DB 481 CTTTGTGGAAGACTTCTTTGTGTAAGAGAGAGTCTCTGCTCGTGTATTTGACCTGGG 540  
QY 1572 GTTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGCTGTCTGGGAAATTTGTG 1631

DB 541 CTTCTCCGGGAATTCAGAAGATGTCGTATGATGCTATTCAGTCTTCTGGGAACTGTGT 600  
QY 1632 CACAAATCACCACACTAGCAGGAACGATGAGTTTTTATATCACCCACCAACTGTCCC 1691  
DB 601 CACAAATCACCACACTAGCAGGAACGATGAGTTTTTATATCACCCACCAACTGTCCC 660  
QY 1692 ATCAGTCTTCGAACTCAACTTCTACAGCAATGGGGTACTTTCATGTTCTTATCATGGAGGC 1751  
DB 661 GTCAAGTCTTCGAACTCAACTTCTACAGCAATGGGGTACTTTCATGTTCTTATCATGGAGGC 720  
QY 1752 CATCATAGCTTGCAGCCTTTATGCAAGTCTCGAA 1784  
DB 721 CATCATAGCTTGCAGCCTTTATGCAAGTCTCGAA 753  
RESULT 10  
AM976326  
LOCUS  
DEFINITION EST388435 MAGE resequenes, MAGN Homo sapiens cDNA, mRNA sequence.  
ACCESSION  
VERSION AM976326.1 GI:8167552  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 580)  
Hagde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt  
J.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and  
Quackenbush,J.  
TITLE  
Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray  
JOURNAL  
COMMENT  
Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johnq@tigr.org  
Plate: 359  
Seq primer: Reverse.  
Location/Qualifiers  
source  
1. 580  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="MAGE resequenes, MAGN"  
/note="Vector: pBluescriptSKm"  
BASE COUNT 174 a 119 c 138 g 149 t  
ORIGIN  
Query Match 22.0%; Score 547.2; DB 10; Length 580;  
Best Local Similarity 97.7%; Pred. No. 8.8e-143;  
Matches 555; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
QY 1889 GCACCATCTCAGTCCTTGGCAGACATTTTACCAAGTCTGCCATGAACACAGTGAAGT 1948  
DB 1 GCTCATCTCAGTCCTTGGCAGACATTTTACCAAGTCTGCCATGAACACAGTGAAGT 60  
QY 1949 TTATCCAGTATGGCATTTTACAGTGGCAGACGATGACCCAGGAAGATATCAGTCTCTA 2008  
DB 61 TTATCCAGTATGGCATTTTACAGTGGCAGACGATGACCCAGGAAGATATCAGTCTCTA 120  
QY 2009 GTCTTGTGACGACGATGGGACAGAGCTTCCAGAACCTTTGTCTTGGAGAGTGTATG 2068  
DB 121 GTCTTGTGACGACGATGGGACAGAGCTTCCAGAACCTTTGTCTTGGAGAGTGTATG 180  
QY 2069 AAGAAGATGAAGCAGTGTCTTGGGAGGACAGCAGATGCTACCTGAGGTGAGCC 2128  
DB 181 AAGAAGATGAAGCAGTGTCTTGGGAGGACAGCAGATGCTACCTGAGGTGAGCC 240  
QY 2129 AATCCAGGAGCACCAGCAGTGTATACCTTCTTACAGAGACTCTCTTGGGCTTTGTCTG 2188

```
Db 241 AATCCAGGAGCACCAGCAGTATACCTTCTTACAGAGCTCCTTTGGSCCTTCTGCTG 300
Qy 2189 AGGCTACAGCTCTGTCGCATCTTTGTTCAACAATTCAGTCTGCTGTTCCAGAACCTG 2248
Db 301 AGGCTACAGCTCTGTCGCATCTTTGTTCAACAATTCAGTCTGCTGTTCCAGAACCTG 360
Qy 2249 AGTATCTGCAAAAGTTGCACAAATACCTAATAACCAAGACAGAAAGAAATGTTGCAAT 2308
Db 361 AGTATCTGCAAAAGTTGCACAAATACCTAATAACCAAGACAGAAAGAAATGTTGCAAT 420
Qy 2309 ATGCTGAGAGTGCACATATTTGCTTGTGAAGATGCTGTGAAGATGTTTAAAGATATTG 2368
Db 421 ATGCTGAGAGTGCACATATTTGCTTGTGAAGATGCTGTGAAGATGTTTAAAGATATTG 480
Qy 2369 GGGTTTTCAAGGAGACCAACCAAGAGAGTGTCTGTTTAGAATGAGCAGCAGCTTTTC 2428
Db 481 GGGTTTTCAAGGAGACCAACCAAGAGAGAGGGCTGTTTAAAGATGAGCAGCAGCTTTT 540
Qy 2429 TACCTCAATGCAACCGACAAACTTCT 2456
Db 541 TACCTAAATGCAACCGACAAACTTCT 568

RESULT 11
LOCUS BG034520 988 bp mRNA linear EST 24-JAN-2001
DEFINITION 602302638F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4404225 5',
mRNA sequence.
ACCESSION BG034520
VERSION BG034520.1 GI:12427920
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM10115 row: k column: 10
High quality sequence stop: 504.
Location/Qualifiers
FEATURES
source
1.988
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4404225"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_87"
/notes="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 269 a 250 c 248 g 221 t
ORIGIN
Query Match 21.3%; Score 530; DB 10; Length 988;
Best Local Similarity 93.8%; Pred. No. 8.1e-138;
Matches 619; Conservative 0; Mismatches 30; Indels 11; Gaps 6;
Qy 1630 GTCAATACCCACACTAGCAGGACGATGATGTTTTTATCACCACCCAGCACAACGTGC 1689
```

```
Db 1 GTCAATACCCACACTAGCAGGATGATGATGTTTTTATCACCACCCAGCACAACGTGC 60
Qy 1690 CCATCAGCTCTCGAATCAACTCTTACAGCAATGGGTACTTTCATGCTTTTATCATGGAG 1749
Db 61 CCATCAGCTCTCGAATCAACTCTTACAGCAATGGGTACTTTCATGCTTTTATCATGGAG 120
Qy 1750 GCATCATAGCTTTGAGCCCTTTATCAGTTCTGAACAAGAGGGGACTGGGGGCTCCACT 1809
Db 121 GCATCATAGCTTTGAGCCCTTTATGAGTTCTGAACAAGAGGGGACTGGGGGCTCCACT 180
Qy 1810 AGCACCCCACTAACTGATCAGCCAGGAGAGCTGGTGGGAAGCGGCCAGCCTGTGC 1869
Db 181 AGCACCCCACTAACTGATCAGCCAGGAGAGCTGGTGGGAAGCGGCCAGCCTGTGC 240
Qy 1870 TACCTTCTCTCAATGAAGGACCATCTCAGTATGGCATTTTACAGTGGCAGCAGATGA 1928
Db 241 TACCTTCTCTCAATGAAGGACCATCTCAGTATGGCATTTTACAGTGGCAGCAGATGA 300
Qy 1929 CCATGAACAGTAGGAAGTTTATCAGTATGGCATTTTACAGTGGCAGCAGATGA 1988
Db 301 CCATGAACAGTAGGAAGTTTATCAGTATGGCATTTTACAGTGGCAGCAGATGA 360
Qy 1989 CCAGGAAGATATCAGTCTTCTGAGCAGCAGTGGGCAAGAAAGCTTCCGAAC 2048
Db 361 CCAGGAAGATATCAGTCTTCTGAGCAGCAGTGGGCAAGAAAGCTTCCGAAC 420
Qy 2049 TTTGCTTGGAGAGTGTGAAGAGATGAACAGCAGTGTGGGGAGGACGCGAGA 2108
Db 421 TTTGCTTGGAGAGTGTGAAGAGATGAACAGCAGTGTGGGGAGGACGCGAGA 480
Qy 2109 TTTGCTTGGAGAGTGTGAAGAGATGAACAGCAGTGTGGGGAGGACGCGAGA 2166
Db 481 TTTGCTTGGAGAGTGTGAAGAGATGAACAGCAGTGTGGGGAGGACGCGAGA 540
Qy 2167 AGA----CTCCTTGGGCTTCTGAGGAGCTTACA--GCTCTGCTGCCATCTTTGTTAC 2220
Db 541 AGAGGAGCTCTTTGGGCTTCTGAGGAGCTTACAAGCTTTGCTGCCATCTTTGTTAC 600
Qy 2221 AACTT--CAGTGTCTGTTCCAGAACCTGAGTATCTGAAAAGTTGCAACATACCTAA 2278
Db 601 AACTTTCAGTGTCTGTTCCAGAACCTTGAATCTGAAAAGTTGCAACATACCTTA 660

RESULT 12
LOCUS CD000106 776 bp mRNA linear EST 01-MAY-2003
DEFINITION AGENCOURT_13642874 NIH_MGC_186 Homo sapiens cDNA clone
IMAGE:30322827 5', mRNA sequence.
ACCESSION CD000106
VERSION CD000106.1 GI:30294625
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDCM124 row: h column: 04
High quality sequence stop: 549.
Location/Qualifiers
FEATURES
source
1..776
```

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30322827"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_186"  
/note="Organ: Pooled Skin; Vector: pDNR-LIB; Site 1: SfiI (ggcattatggcc); Site 2: SfiI (ggcgccctggcc); Library is oligo-dT primed and directionally cloned. cDNA was prepared from a pooled samples of tissues from Skin, meninges, duramater, pia matter and choroid plexus. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATCTAGAGCGGAGCGGCACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library"  
BASE COUNT 208 a 180 c 199 g 189 t  
ORIGIN  
Query Match 20.6%; Score 511.4; DB 14; Length 776;  
Best Local Similarity 99.8%; Pred. No. 1.3e-132;  
Matches 512; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1972 GTGGCAGACGACGATGACCGAAGATATCAGTCTCTAGTCTGTCTGCTGAGCAGAGTGGGAC 2031  
DB 1 GGGCAGAGCAGGATGACCGAGAGATATCAGTCTCTAGTCTGTCTGAGCAGAGTGGGAC 60  
QY 2032 AAGAAGCTTCAGAACCTTTGTCTGGAGAAGTGATGAAGAAGATGAAGACAGTACTTT 2091  
DB 61 AAGAAGCTTCAGAACCTTTGTCTGGAGAAGTGATGAAGAAGATGAAGACAGTACTTT 120  
QY 2092 GGGAGGAGACCGAGATGCTACTGAGGTGAGCCTAATCGAGGAGCACAGCAGTTT 2151  
DB 121 GGGAGGAGAACGCGAGATTGTACTGGAAGGTGAGCCTAATCGAGGAGCACAGCAGTTT 180  
QY 2152 ATCACCCTTTACAGAGACTCTTGGGCTTTGTCTGGAGGCTACAGCTCTGCTGCCATC 2211  
DB 181 ATCACCCTTTACAGAGACTCTTGGGCTTTGTCTGGAGGCTACAGCTCTGCTGCCATC 240  
QY 2212 TTGTTCACACTTCAGTGGTCTCTGTTCCAGAACCTGAGTATCTGCMAAAGTTGCACAAA 2271  
DB 241 TTGTTCACACTTCAGTGGTCTCTGTTCCAGAACCTGAGTATCTGCMAAAGTTGCACAAA 300  
QY 2272 TACCTAATAACAGAACGAAGAATGTTGAGATATGCTGAGAGTGCACATATGTT 2331  
DB 301 TACCTAATAACAGAACGAAGAATGTTGAGATATGCTGAGAGTGCACATATGTT 360  
QY 2332 CTTGTGAAGATGCTGTGAAAATGTTTAAAGTATTTGGGTTTTCAGGAGACCAACAA 2391  
DB 361 CTTGTGAAGATGCTGTGAAAATGTTTAAAGTATTTGGGTTTTCAGGAGACCAACAA 420  
QY 2392 AAGAGAGTGTCTGTTTAAAGTCTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAA 2451  
DB 421 AAGAGAGTGTCTGTTTAAAGTCTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAA 480  
QY 2452 CTTCTAGAATATATCTGAGTTTGTGTGTG 2484  
DB 481 CTTCTAGAATATATCTGAGTTTGTGTGTG 513

RESULT 13  
BQ694822  
LOCUS  
DEFINITION BQ694822 550 bp mRNA linear EST 15-JUL-2002  
BO694822 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens  
cDNA 5', mRNA sequence.  
ACCESSION BQ694822  
VERSION BQ694822.1 GI:21820138  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (Bases 1 to 550)  
AUTHORS Yang, R.-Z., Shuldiner, A. and Gong, D.-W.  
TITLE EST analysis of human adipose gene expression  
JOURNAL Unpublished  
COMMENT Contact: Gong Da-Wei  
Division of Endocrinology, Diabetes and Nutrition  
University of Maryland  
660 Redwood St, HH497, Baltimore, MD 21201, USA  
Tel: 410 706 1672  
Fax: 410 706 1622  
Email: dgong@medicine.umaryland.edu  
PCR Primers  
FORWARD: CTCGGGAGCGCGCCATTGTCTTGGT  
BACKWARD: AATACGACTCACTATAGGCGGAATTGG  
Seq primer: GTTGTACCCCGGAATTC.  
Location/Qualifiers  
FEATURES  
1..550  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/sex="Male and Female"  
/tissue\_type="Adipose"  
/clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"  
/note="Vector: lambdaTriplex"  
BASE COUNT 145 a 140 c 126 g 136 t 3 others  
ORIGIN  
Query Match 20.0%; Score 496.8; DB 13; Length 550;  
Best Local Similarity 98.5%; Pred. No. 1.5e-128;  
Matches 542; Conservative 0; Mismatches 4; Indels 4; Gaps 4;  
QY 1454 CACACATTTGGCTTGGCTGCTCTCTACAGACAGCAGGAGGAATTGATCTCTCCACAT 1513  
DB 2 CACACATTTGGCTTGGCTGCTCTCTACAGACAGCAGGAGGAATTGATCTCTCCACAT 61  
QY 1514 TGGTCGAAGACTCTTTGTGATGAAGAGGAAGTCTGGCTCGTGAATTTTACCTGGGGT 1573  
DB 62 TGGTCGAAGACTCTTTGTGATGAAGAGGAAGTCTGGCTCGTGAATTTTACCTGGGGT 121  
QY 1574 TCTCAGGAATTCAGAAAGATGTAGTAATGATGCATCCATACAGCTCTCGGAAATTTGTGCA 1633  
DB 122 TCTCAGGAATTCAGAAAGATGTAGTAATGATGCATCCATACAGCTCTCGGAAATTTGTGCA 181  
QY 1634 CAATCACCACACATAGCAGGAACGATGAGTTTTTTTATCACCACAGCAAACTGTCCTAT 1693  
DB 182 CAATCACCACACATAGCAGGAATGATGAGTTTTTTTATCACCACAGCAAACTGTCCTAT 240  
QY 1694 CAGTCTTGGAACTCAACTTCTACAGCAATGGGTTACTTCTATGCTTTTATCATGAGAGCCA 1753  
DB 241 CAGTCTTGGAACTCAACTTCTACAGCAATGGGTTACTTCTATGCTTTTATCATGAGAGCCA 300  
QY 1754 TCATAGCTTG-CAGCCTTTTATGAGTTCTGAAAGAGGAGTGGGGTCCCACTAGC 1812  
DB 301 TCATAGCTTGAGCCTTTTATGAGTTCTGAAAGAGGAGTGGGGTCCCACTAGC 360  
QY 1813 ACCCCACCTAACCTGATC-AGCCAGGAGCAGCTGGTGGG-AAGGCGGCCAGCCTGTGCT 1870  
DB 361 ACCCCACCTAACCTGATCAAGCCAGGAACAGTGGTGGGAAAGGCGGCCAGCCTGTGCT 420  
QY 1871 ACCTTCTCTCAATGAAGGACCACTCTCACTGCTTGGCCAGACATTTTACCAAGTCTGCC 1930  
DB 421 ACCTTCTCTCAATGAAGGACCACTCTCACTGCTTGGCCAGACATTTTACCAAGTCTGCC 480  
QY 1931 ATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTCTTACAGTGGCAGAGCAGCATGACC 1990  
DB 481 ATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTCTTACAGTGGCAGAGCAGCATGACC 540  
QY 1991 AGGAAGATAT 2000  
DB 541 AGGAAGATAT 550

```
RESULT 14
BI691948          690 bp  mRNA  linear  EST 18-SEP-2001
LOCUS             603311882F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5351781 5',
DEFINITION        mRNA sequence.
ACCESSION         BI691948
VERSION           BI691948
KEYWORDS          EST.
SOURCE            Mus musculus (house mouse)
ORGANISM          Mus musculus
REFERENCE         1 (bases 1 to 690)
AUTHORS          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE            NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL          National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT          Unpublished
                  Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-i@mail.nih.gov
                  Tissue Procurement: Jeffrey Green M.D.
                  CDNA Library Preparation: Life Technologies, Inc.
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LNL at:
                  http://image.lnl.gov
                  Plate: L1AM11894 Row: d Column: 22
                  High quality sequence stop: 687.
FEATURES
    source
        1. 690
            Location/Qualifiers
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:5351781"
                /sex="female, virgin"
                /tissue_type="infiltrating ductal carcinoma"
                /dev_stage="5 months"
                /lab_host="DH10B"
                /clone_lib="NCI CGAP Mam6"
                /notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
                Library constructed by Life Technologies. Investigator
                providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      173 a 172 c 180 g 165 t
ORIGIN
    Query Match      19.6%; Score 487.6; DB 12; Length 690;
    Best Local Similarity 85.5%; Pred. No. 6.5e-126;
    Matches 590; Conservative 0; Mismatches 94; Indels 6; Gaps 4;
QY 1493 AGGAATTGATCTCTCCACATTTGGTGGAACTCTTTGTGATGAAAGAGAAAGTCTCTGG 1552
Db 1 AGGGAATCCATCTCTCCACGCT-GTGGAACTCTTTGTGATGAAAGAGAAAGTCTCTAG 59
QY 1553 CTGCTGATTTTGACCTGGGGTTCTCAGGAATTCAGAGATGATAGTATCATGCCATAC 1612
Db 60 CTCGCGATTTTCGACCTAGGCTTCTCGGGAAATTCAGAAAGATTCGTCATGCTATTC 119
QY 1613 AGCTGCTGGGAAATTTGTGCACATACCCACACTAGCAGAACGATGAGTTTTTATCA 1672
Db 120 AGCTTCTGGGAACTGTGTACATATCCACACGAGCAGGAAGATGAGTTTTTATTA 179
QY 1673 CCCCCAGCACAACGTGTCCTCCATCAGTCTTCGAACTCAAC-TTCTACAGCAATGGGTACTT 1731
Db 180 CTCGCCAGCACAACGTGCTCCGCTCAGTCTTTGAACTCAACGCTTCTACAGCAATGGGTACTT 239
QY 1732 CATGTCTTTATCATGAGGCGCATCATAGCTTGAGCCTTTATGAGTCTTGACAGAGG 1791
Db 240 CATGTGTTTCATGGAAGCCATCATAGCTTGAGCATCTATGAGTCTTGAATAGAGG 299
QY 1792 GGACTGGGGGGTCCCACTAGCACCCACCTAACTGATCAGCCAGGAGCAGCTGGTGGCG 1851
Db 300 TGCTCTGGAGGGTCCGCTCGAGGCTCGGCAACCTGATCAGCAGGAGCAGCTGGTGAGG 359
QY 1852 AAGGGCGGCAGCCTGTGCTACCTTCTCTCAATGAGGCAACCATCTCTACTGCTTGCAG 1911
Db 360 AAGGGCGGCAGCCTGTGCTACCTTCTCTCAATGAGGCAACCATCTCTACTGCTTGCAG 419
QY 1912 ACATTTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTCTTACA 1971
Db 420 ACTTTTACCAAGTCTGTCTATGAGACAGTTGGCAAGTTTCATCCAGTATGGCATTTCTACA 479
QY 1972 GTGGCAGACGACGATGACCCAGGAAGATATCATGCTCTAGTCTTGTGCTGAGCAGAGTGGAC 2031
Db 480 GTGGCAGACGAAAGATGACCCAGGAAGATGTCAGTCTCTGGCTTGCAGAACAGCAGTGGAC 539
QY 2032 AAGAGCTTCCAGAACCTTTGTCTTGGAGAACTGATGAAGAGATGAACACAGTGAATTT 2091
Db 540 AAGAAGCTTCCGGAAC---TGAAGTGGAGAACTGAGGAAAGATGAACACAGTGAATTT 596
QY 2092 GGGGAGGA-ACAGCGAGATTGCTTACTGAAGGTGAGCCCAATCAAGGAGCAGCAGCAGTT 2150
Db 597 GGTGAGGAGCCAGCGAGATTGCTTACTCAAGGTGAGCCAGTCCAGGAGCAGCAGCAATT 656
QY 2151 TATCAGCTTCTTACAGAGACTCTCTTGGGCC 2180
Db 657 CATCAGCTTCTTACAGAGCTTCTAGGTCC 686
RESULT 15
BU263739          825 bp  mRNA  linear  EST 26-NOV-2002
LOCUS             60381507F1 CSEQCHN52 Gallus gallus cDNA clone CHEST808n1 5', mRNA
DEFINITION        sequence.
ACCESSION         BU263739
VERSION           BU263739.1 GI:255333451
KEYWORDS          EST.
SOURCE            Gallus gallus (chicken)
ORGANISM          Gallus gallus
REFERENCE         1 (bases 1 to 825)
AUTHORS          Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
                  Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
                  A Comprehensive Collection of Chicken cDNAs
TITLE            Curr. Biol. 12 (22), 1965-1969 (2002)
JOURNAL          MEDLINE
PUBMED           22335534
COMMENT          Contact: Simon Hubbard
                  Department of Biomolecular Sciences
                  University of Manchester Institute of Science and Technology (UMIST)
                  PO Box 88, Manchester, M60 1QD, UK
                  Tel: 01612008930
                  Fax: 01612360409
                  Email: Simon.Hubbard@umist.ac.uk.
FEATURES
    source
        1. 825
            Location/Qualifiers
                /organism="Gallus gallus"
                /mol_type="mRNA"
                /strain="White Leghorn, Hisex"
                /db_xref="taxon:9031"
                /clone="CHEST808n1"
                /dev_stage="22"
                /lab_host="DH10B"
                /clone_lib="CSEQCHN52"
                /note="Organ: limbs; Vector: pBluescript II KS(+); Site 1:
                EcoRI; Site 2: NotI; This normalized library was
                constructed from 1 million independent clones. cDNA
                synthesis was initiated using an oligo(dT) primer, using
                methylated C in the first strand synthesis reaction.
                Following this first strand reaction, double-stranded cDNA
                was blunted, ligated to NotI adapters, digested with EcoRI
```

, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 231 a 178 c 182 g 234 t  
ORIGIN

Query Match 18.8%; Score 467.8; DB 13; Length 825;  
Best Local Similarity 74.5%; Pred. No. 2.8e-120;  
Matches 615; Conservative 0; Mismatches 207; Indels 3; Gaps 2;

QY 1001 TTGTGGTAGATACTCTGTCTACCAATGTCACCCAGACATCTTGATATACCTGTGTGAA 1060  
DB 1 TGGTAGTGGATGCTCTCTGTGCAATGCTACTCTCTGATGCTCTAATATACCTGTGGAA 60

QY 1061 TCTCTATATGCGATATATCGAAGTCACTACATGGTGAACAACCTGGGCAAACTTAAGA 1120  
DB 61 TCTCTACGATCGCATAAATTGAAGTCACTATAACAGTGAACAGCTGGGCAAGCTTAAGA 120

QY 1121 AGAATGACAGCCTGTGGAGTGTAGCAAGAGTGTATTAGAAATGTTACGAAAAACTATG 1180  
DB 121 AGAATGAAGTCTTTGGAGTATAGCAAGAGAGTCTTCAGAAATGCTGGGAAGAAATTATG 180

QY 1181 GTTGTGTCGGAGTGGATTGTCACAGCCATTTCTTAAAGAAATATTAGAAAGCCAAA 1240  
DB 181 GGTGTGTCAGATAGATTGTCACAAACATTTCTTAAAGAAATATGTGAACGCCAAA 240

QY 1241 GTCAGAAACCGGTGTCCTCTACTTTCCCTGGAGCAAGCGTGTACCAGCTATACCTTC 1300  
DB 241 GTCAGAAACCGTCTCTCTCTCTCTTCTTGAACAAGCTTTGTTACCAGCTATACCTTC 300

QY 1301 CTTCAAGACCCAGTATGCTGCTGATGAAGTAGAGACACGTCCTCAATTAATGAGTCCAGAA 1360  
DB 301 CATCGAGACCTAAATGATATGTGGATGAAGTACTGAGGCTTCACTGCTCAACTCCAAG 360

QY 1361 ATGCAACAGATGAATCCCTACGAGGAGTGTGATTGCAAAATCTGGCTGAGCATATTTAT 1420  
DB 361 ATATAACAGTGAACCCCTACAGAGAGAGCTGTAGCCAAATTTGGCTGAGCATATTTAT 420

QY 1421 TCATGTGTAGCAAGTCTGTGCCATATATGTCCACACATTTGGCTTGCCTCTCTCT 1480  
DB 421 TCATGTGTAAACAAGTCTGTGTGTGTCTACCCACATTTGTCCTGTTTGTGCTGT 480

QY 1481 ACAGACACAGGAGGAAATGATCTCTCCATTTGGTGAAGACATTTCTTTGTGATGAAG 1540  
DB 481 ACAGACATAGACAGGGGACTGATCTTTCCAGGTTGGTGAAGATTTCTTCTCCATGAAG 540

QY 1541 AGGAAGTCTGCTGATTTTGAACCTGGGGTCTCAGGAAATTCAGAAATGTAGTAA 1600  
DB 541 AAGAGTCTTAGCCGCTGACTTTCGACTTGGAAATTTTCAGGAACTCAGATGATTTGCA 600

QY 1601 TGCATGCCATACAGTCTGTGGAAATTTGTCACAATCACCCACACTAGCAGGAACCATG 1660  
DB 601 TGCATGCCATCCACTTGTGTGGGAACTGTGTAATATCAACAACACTAGTGAACAACG 660

QY 1661 AGTTTTTTTAT -CACCCACAGACAACTGTCCCATCAGTCTTCGAACTCAACTTCTACAGC 1719  
DB 661 AGTTCTTCATAAACTCCACAGACACGATCCCTGCTGCTTTTGAACCTCAATTTCTACAGC 720

QY 1720 AATGGGTACTTTCATGCTTTTATCATGAGGCCATCATAGCTTGA -GCCTTTATGAG 1777  
DB 721 AATGGAATCTTCATGATTCATTAAAGAGGAGTCAATGGGCTTCAAGTCTTCATGCCAG 780

QY 1778 TTCTGAACAAGAGGGGAGTGGGGGTCCCACTAGCACCCCACTTA 1822  
DB 781 TTCAGATGAAGGTCCCGAAATGGTACCAATGGTGTCTTCTCCCA 825

THIS PAGE BLANK (USPTO)

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2484	100.0	3003	24	ABK11094	CDNA encoding huma
2	2484	100.0	4031	23	ABV25313	Human prostate exp
3	2482.4	99.9	2487	24	ABK94821	DNA encoding huma
4	2480.8	99.9	2487	25	AAL55475	GPAM related DNA s
5	2480.8	99.9	3210	24	AD352221	Human TRNF-6 cDNA
6	2476	99.7	2684	25	ABX72233	Human NOVX polynuc
7	2185.8	88.0	3273	24	ABZ12025	Human polynucleoti
8	1879.2	75.7	2646	24	ABK63741	Rat sequence diffe

9	437	17.6	469	22	ABA08415	Human sn-glycerol-
10	342.2	13.6	429	25	ABX42219	Bovine EST associa
11	337.6	13.6	407	25	ABX43656	Bovine EST associa
12	302	12.2	418	25	ABX35577	Bovine EST associa
13	233	9.4	303	25	ABX47921	Bovine EST associa
14	209.6	8.4	264	25	ABX41074	Bovine EST associa
15	172.6	6.9	387	25	ABX42469	Bovine EST associa
16	161.8	6.5	231	24	ABN22191	Human ORFX polynuc
17	109	4.4	2682	24	ABSS52820	CDNA encoding huma
18	109	4.4	2755	24	ABD35222	Human TRNFR-7 cDNA
19	103.8	4.2	3439	23	ABL09555	Drosophila melanog
20	103.8	4.2	9676	23	ABL09554	Drosophila melanog
21	99.4	4.0	501	25	ABX38896	Bovine EST associa
22	53.8	2.2	2640	23	ABSR87279	DNA encoding novel
23	49.2	2.0	65	24	ABN28887	Rat spliced transcr
24	47	1.9	2502	22	AD020332	E. coli p1eB DNA e
25	47	1.9	2502	22	AD020333	Escherichia coli p
26	46.2	1.9	2154	23	ABSR81996	DNA encoding novel
27	46.2	1.9	2154	23	ABSR82407	DNA encoding novel
28	46.2	1.9	2154	23	ABSR82517	DNA encoding novel
29	45.6	1.8	2175	23	ABL07919	Drosophila melanog
30	45	1.8	100848	22	ABF28552	Genomic fragment #
31	44.8	1.8	1830121	17	ABT42063	Haemophilus influe
32	43.6	1.8	1127	23	ABSR81997	DNA encoding novel
33	43.6	1.8	1127	23	ABSR82408	DNA encoding novel
34	43.6	1.8	4120	23	ABSR90070	DNA encoding novel
35	41.2	1.7	185371	24	ABT10718	Human breast canc
36	39.8	1.6	21045	22	ABSR26721	Human genomic DNA
37	39.8	1.6	21045	25	ABX74070	Human novel polynu
38	39	1.6	1673	23	ABSR72799	DNA encoding novel
39	39	1.6	2114	21	ABSR74225	Human secreted pro
40	39	1.6	3714	23	ABSR73941	DNA encoding novel
41	38.6	1.6	4379	22	ABSR2496	Human immune/haema
42	38.6	1.6	4794	22	ABSR82497	Human immune/haema
43	38.6	1.6	4799	22	ABSR82500	Human immune/haema
44	38.6	1.6	4807	22	ABSR2498	Human immune/haema
45	38.6	1.6	1664976	19	AAV21209	Methanococcus jann

## RESULT 1

ABK11094

ID ABK11094 standard: cDNA: 3003 bp.

AC ABK11094;

DT 18-JUN-2002 (first entry)

DE cdNA encoding human acyltransferase. ACTR-1.

Human; acyltransferase; ACAT-1; antidiabetic; antilipaeamic; anorectic;  
antihypertensive; cerebroprotective; anorectic; ophthalmological;  
cardiac; metabolic disorder; energy homeostasis disorder; diabetes;  
hyperglycaemia; hypercholesterolaemia; hyperlipoproteinaemia; stroke;  
hypertriacylglyceridaemia; hyperlipidaemia; atherosclerosis; obesity;  
retinopathy; nephropathy; peripheral neuropathy; weight disorder;  
appetite regulation disorder; cachexia; anorexia; bulimia;  
cardiovascular disorder; gene; ss.

Homo sapiens.

XX	Key	Location/Qualifiers
FH		

FT	key	Education/
FT	CDS	341..2827

FT  
/\*tag= a

```

FT      /product= "Human acyltransferase ACPR-1"
FT      /note= "The coding region is specifically claimed
FT      in claim 1"

```

PN WO200216592-A2

28-FEB-2002





Db 1601 CTACTTTCCCTGGAGCAACGGTTGTTACGAGCTATCTCTTCAAGACCAGATGCT 1660  
Qy 1321 GCTGATGAAGTAGAGACACGTCCTAATTAATGAGTCCAGAAATGCAACAGATGCAATCCCTA 1380  
Db 1661 GCTGATGAAGTAGAGACACGTCCTAATTAATGAGTCCAGAAATGCAACAGATGCAATCCCTA 1720  
Qy 1381 CGAAGGAGGTTCATTGCAAAATCTGCTGAGCATATTTCTATTACCTGCTAGCAAGTCTCTGT 1440  
Db 1721 CGAAGGAGGTTCATTGCAAAATCTGCTGAGCATATTTCTATTACCTGCTAGCAAGTCTCTGT 1780  
Qy 1441 GCATTATGTCCACACACATTTGCTGCTGCTGCTGCTCTCTACAGACACAGCAGGGAAT 1500  
Db 1781 GCATTATGTCCACACACATTTGCTGCTGCTGCTGCTCTCTACAGACACAGCAGGGAAT 1840  
Qy 1501 GATCTCTCCACATTTGCTGCAAGACTTTCTTTGTGATGAAGAGGAAGTCTGCTGCTGAT 1560  
Db 1841 GATCTCTCCACATTTGCTGCAAGACTTTCTTTGTGATGAAGAGGAAGTCTGCTGCTGAT 1900  
Qy 1561 TTTGACCTGGGGTTCTCAGGAATTTCAAGAGTGTAGTAATGCATGCCATACAGCTGCTG 1620  
Db 1901 TTTGACCTGGGGTTCTCAGGAATTTCAAGAGTGTAGTAATGCATGCCATACAGCTGCTG 1960  
Qy 1621 GGAATTTGTGTCCAAATCACCACACTAGCAGGAACGATGATTTTTTTATCACCCCCAGC 1680  
Db 1961 GGAATTTGTGTCCAAATCACCACACTAGCAGGAACGATGATTTTTTTATCACCCCCAGC 2020  
Qy 1681 ACACTGTCCCATCAGTCTTCCAACTCTACAGCAATGGGGTACTTCTATCTCTTT 1740  
Db 2021 ACACTGTCCCATCAGTCTTCCAACTCTACAGCAATGGGGTACTTCTATCTCTTT 2080  
Qy 1741 ATCATGAGGGCCATCATAGCTTGGAGCCTTTATGCAAGTCTTGAACAAGAGGGAATGCGG 1800  
Db 2081 ATCATGAGGGCCATCATAGCTTGGAGCCTTTATGCAAGTCTTGAACAAGAGGGAATGCGG 2140  
Qy 1801 GGTCCCACTAGCACCCCACTAACCCTGATCAGCAGGAGCAGCTGGTGGGAGGCGGCC 1860  
Db 2141 GGTCCCACTAGCACCCCACTAACCCTGATCAGCAGGAGCAGCTGGTGGGAGGCGGCC 2200  
Qy 1861 AGCTGTGCTACTCTCTCCAAATGAAGGCAACCATCTCACTGCCCTTGCAGACATTTTAC 1920  
Db 2201 AGCTGTGCTACTCTCTCCAAATGAAGGCAACCATCTCACTGCCCTTGCAGACATTTTAC 2260  
Qy 1921 CAACTGTGCTGATGAACAGTAGAGAAATTTATCCAGTATGGCAATTTTACAGTGCCAGAG 1980  
Db 2261 CAACTGTGCTGATGAACAGTAGAGAAATTTATCCAGTATGGCAATTTTACAGTGCCAGAG 2320  
Qy 1981 CACGATGACCAAGAGATATCAGTCTCTGCTGAGCAGCAGTGGGACAGAGCTT 2040  
Db 2321 CACGATGACCAAGAGATATCAGTCTCTGCTGAGCAGCAGTGGGACAGAGCTT 2380  
Qy 2041 CCAGAACCTTTGCTTTGGAGAGTGCATGAAGAAGATGAAGCAGTGACTTTTGGGGAGGAA 2100  
Db 2381 CCAGAACCTTTGCTTTGGAGAGTGCATGAAGAAGATGAAGCAGTGACTTTTGGGGAGGAA 2440  
Qy 2101 CAGCAGATTTGCTTACCTGAAGTGAAGCAATCCAAAGGAGCAGCAGCTTATCACCTTC 2160  
Db 2441 CAGCAGATTTGCTTACCTGAAGTGAAGCAATCCAAAGGAGCAGCAGCTTATCACCTTC 2500  
Qy 2161 TTACAGACCTCTTTGGGCTTTGCTGAGGCTTACAGCTGCTGCTGCAATCTTTGTTTAC 2220  
Db 2501 TTACAGACCTCTTTGGGCTTTGCTGAGGCTTACAGCTGCTGCTGCAATCTTTGTTTAC 2560  
Qy 2221 AACTTCAGTGTCTGTTCCAGAACCTCAGTATCTGCAAAAGTTCGCAAAATACCTTAATA 2280  
Db 2561 AACTTCAGTGTCTGTTCCAGAACCTCAGTATCTGCAAAAGTTCGCAAAATACCTTAATA 2620  
Qy 2281 ACCAGACAGAAAGAAATTTGTCAGTATATGCTGAGAGTGGCACAATATGCTTGTGAAG 2340  
Db 2621 ACCAGACAGAAAGAAATTTGTCAGTATATGCTGAGAGTGGCACAATATGCTTGTGAAG 2680  
Qy 2341 AATGCTGTGAAATTTTAAAGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGGTG 2400  
Db 2681 AATGCTGTGAAATTTTAAAGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGGTG 2740

Qy 2401 TCTGTTTGAACCTGAGCAGCACTTTTCTACCTCAATGCAACCGCAAAACCTTCTAGAA 2460  
Db 2741 TCTGTTTGAACCTGAGCAGCACTTTTCTACCTCAATGCAACCGCAAAACCTTCTAGAA 2800  
Qy 2461 TATATTCGAGTTTGTGGTCTG 2484  
Db 2801 TATATTCGAGTTTGTGGTCTG 2824

## RESULT 2

ABV25313  
ID ABV25313 standard; cDNA; 4031 BP.

XX AC ABV25313;  
XX 16-SEP-2002 (first entry)  
XX Human prostate expression marker cDNA 25304.  
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX Human; pharmacogenomic marker; gene; ss.  
XX Homo sapiens.

OS WO200160860-A2.  
XX 23-AUG-2001.  
XX 20-FEB-2001; 2001WO-US05171.  
XX 17-FEB-2000; 2000US-183319P.  
XX 16-MAR-2000; 2000US-189862P.  
XX 25-MAY-2000; 2000US-207454P.  
XX 09-JUN-2000; 2000US-211314P.  
XX 18-JUL-2000; 2000US-219007P.  
XX 13-DEC-2000; 2000US-255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -

Claim 1; Page 4980-4981; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:

(a) assessing whether a patient is afflicted with prostate cancer;

(b) monitoring the progression of prostate cancer in a patient;

(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;

(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;

(e) selecting a composition for inhibiting prostate cancer in a patient;

(f) assessing the prostate cell carcinogenic potential of a compound;

(g) determining whether prostate cancer has metastasized in a patient;

(h) assessing the aggressiveness or indolence of prostate cancer in a patient;

(i) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 4031-BP; 1105 A; 882 C; 918 G; 1120 T; 6 other;

Query Match 100.0%; Score 2484; DB 23; Length 4031;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGATGAATCTGCATGACCTTGGTACATAAGATGTTTCTTATCTGCCACATTCATCA 60

Db 341 ATGGATGAATCTGCACTGACCCCTGGTACAAATAGATGTTCTTATCTGCCACATTCATCA 400  
Qy 61 GAATACAGTGTGGTGGATGATGACACACAAAGTGGGATGGGGTGGATGGCTTTAGA 120  
Db 401 GAATACAGTGTGGTGGATGATGACACACAAAGTGGGATGGGGTGGATGGCTTTAGA 460  
Qy 121 CCCACCGTCTTCAGATCTGCAACTTTTAAATGGAAGAAAGCCTTAATAGTTCGGAAGG 180  
Db 461 CCCACCGTCTTCAGATCTGCAACTTTTAAATGGAAGAAAGCCTTAATAGTTCGGAAGG 520  
Qy 181 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCAGAGCTGGGACAAATTTTCAAC 240  
Db 521 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCAGAGCTGGGACAAATTTTCAAC 580  
Qy 241 CCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAGTCAACAAGA 300  
Db 581 CCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAGTCAACAAGA 640  
Qy 301 CACCGGGATGGCTTGCAAGACGCTTTCTTACGTTCTTTTATCAAGAGCGAGATGTG 360  
Db 641 CACCGGGATGGCTTGCAAGACGCTTTCTTACGTTCTTTTATCAAGAGCGAGATGTG 700  
Qy 361 CATAAGGGCATGTTTGCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 420  
Db 701 CATAAGGGCATGTTTGCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 760  
Qy 421 GAGGCAATTCAGAGTGGCTGTGAAATTAACCCCTGATGGTCTGCCACAGCAATCA 480  
Db 761 GAGGCAATTCAGAGTGGCTGTGAAATTAACCCCTGATGGTCTGCCACAGCAATCA 820  
Qy 481 AAGCGCTTAAACAGTGAAGAAAGAGCTTAAAGGATCTTCAAGAAATGGTTGCCACT 540  
Db 821 AAGCGCTTAAACAGTGAAGAAAGAGCTTAAAGGATCTTCAAGAAATGGTTGCCACT 880  
Qy 541 GTCTACCGGCAATGATCAGACTGACTGGGTGGTGTCTGCTAAACCTGTTCAACAGCTTC 600  
Db 881 GTCTACCGGCAATGATCAGACTGACTGGGTGGTGTCTGCTAAACCTGTTCAACAGCTTC 940  
Qy 601 TTTTGAACATTCAAATTCACAAAGTCAAAGTCAAAGTCAAAGTCAAAGTCAAAGTCAAAG 660  
Db 941 TTTTGAACATTCAAATTCACAAAGTCAAAGTCAAAGTCAAAGTCAAAGTCAAAGTCAAAG 1000  
Qy 661 AATTTGCGCTCTGTTTCTACAGTTCATAGATCCCATATTCAGTATCTGCTGCTCACT 720  
Db 1001 AATTTGCGCTCTGTTTCTACAGTTCATAGATCCCATATTCAGTATCTGCTGCTCACT 1060  
Qy 721 TTCATTTCTTCTGCCATTAACATCAAGCAACCATACATTTGCTTTCAGGCAATATCTCAAC 780  
Db 1061 TTCATTTCTTCTGCCATTAACATCAAGCAACCATACATTTGCTTTCAGGCAATATCTCAAC 1120  
Qy 781 ATCCCAATCTTCAGTACCTTGATCCATAAGCTGGGGCTTCTTTCATACGACGAAGCTC 840  
Db 1121 ATCCCAATCTTCAGTACCTTGATCCATAAGCTGGGGCTTCTTTCATACGACGAAGCTC 1180  
Qy 841 GATGAACACCAAGTGGAGGAAGTGTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db 1181 GATGAACACCAAGTGGAGGAAGTGTCTCTATAGAGCTTTGCTCCATGGGCATATA 1240  
Qy 901 GTTGAATTTATTCGACAGCAGCAATTTCTTGGAGATCTTCTGGAAGCACAACCTTCTAGG 960  
Db 1241 GTTGAATTTATTCGACAGCAGCAATTTCTTGGAGATCTTCTGGAAGCACAACCTTCTAGG 1300  
Qy 961 AGTGGAAAAACCTTTGTGCTCGGCAAGACTTTTGTGTCAGTGTGGTAGATCTGCTGTCT 1020  
Db 1301 AGTGGAAAAACCTTTGTGCTCGGCAAGACTTTTGTGTCAGTGTGGTAGATCTGCTGTCT 1360  
Qy 1021 ACCAATGTCAATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATCCGATATC 1080  
Db 1361 ACCAATGTCAATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATCCGATATC 1420  
Qy 1081 GAAGGTCACTACAATGGTGAACAACTGGGCAAACTTAAGAAAGATGAGAGCTTGTGGAGT 1140

Db 1421 GAAGGTCACTACAATGGTGAACAACTGGGCAAACTTAAGAAAGATGAGAGCTTGTGGAGT 1480  
Qy 1141 GTAGCAAGAGTGTATTATAGAACTTACGAAAACTATGGTGTGTCCGAGTGGATTTT 1200  
Db 1481 GTAGCAAGAGTGTATTATAGAACTTACGAAAACTATGGTGTGTCCGAGTGGATTTT 1540  
Qy 1201 GCACAGCCATTTTCTTAAAGGAATATTATTAGAAAGCAAAGTCAAGAACCGGTGTCTGCT 1260  
Db 1541 GCACAGCCATTTTCTTAAAGGAATATTATTAGAAAGCAAAGTCAAGAACCGGTGTCTGCT 1600  
Qy 1261 CTACTTTCCCTGGAGCAAGCGTTGTTAACAAGTATCTTCTTCAAGACCCAGTGTGCT 1320  
Db 1601 CTACTTTCCCTGGAGCAAGCGTTGTTAACAAGTATCTTCTTCAAGACCCAGTGTGCT 1660  
Qy 1321 GCTGATGAAGTGAAGACACAGTCCATTAATAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db 1661 GCTGATGAAGTGAAGACACAGTCCATTAATAGTCCAGAAATGCAACAGATGAATCCCTA 1720  
Qy 1381 CGAAGGAGTTCATTTGCAAAATCTGGCTGAGCATATTTCTATTCACTGCTAGCAAGTCTCT 1440  
Db 1721 CGAAGGAGTTCATTTGCAAAATCTGGCTGAGCATATTTCTATTCACTGCTAGCAAGTCTCT 1780  
Qy 1441 GCCATTAATGTCACACATTTGCTGCTGCTGCTCTCTAAGACACAGGCAAGGAAT 1500  
Db 1781 GCCATTAATGTCACACATTTGCTGCTGCTGCTCTCTAAGACACAGGCAAGGAAT 1840  
Qy 1501 GATCTCTCACATTTGGTGAAGACTTCTTTGATGAAGAGGAGTCTGCTGCTGAT 1560  
Db 1841 GATCTCTCACATTTGGTGAAGACTTCTTTGATGAAGAGGAGTCTGCTGCTGAT 1900  
Qy 1561 TTTTGACCTGGGCTTCTCAGAAATTCAGAAAGTGTAGTAATGCATGCCATACAGCTGCTG 1620  
Db 1901 TTTTGACCTGGGCTTCTCAGAAATTCAGAAAGTGTAGTAATGCATGCCATACAGCTGCTG 1960  
Qy 1621 GGAATTTGTCACAAATCAACCACTAGCAGGAACGATGAGTGTATATCACCCACG 1680  
Db 1961 GGAATTTGTCACAAATCAACCACTAGCAGGAACGATGAGTGTATATCACCCACG 2020  
Qy 1681 ACAATGTCCTCATGCTTCTGAACTCACTTCTACAGCAATGGGTACTCATGCTTTT 1740  
Db 2021 ACAATGTCCTCATGCTTCTGAACTCACTTCTACAGCAATGGGTACTCATGCTTTT 2080  
Qy 1741 ATCATGGAGGCCATCATAGCTTTCAGCCCTTTATGCAAGTCTTCTGAAAGAGGAGTCTGGG 1800  
Db 2081 ATCATGGAGGCCATCATAGCTTTCAGCCCTTTATGCAAGTCTTCTGAAAGAGGAGTCTGGG 2140  
Qy 1801 GGTCCCACTAGCACCCCACTAACTGATCAGCAGGAGCAGTGTGCGGAAGGCGGC 1860  
Db 2141 GGTCCCACTAGCACCCCACTAACTGATCAGCAGGAGCAGTGTGCGGAAGGCGGC 2200  
Qy 1861 AGCCTGTGTACTTCTCTCCAATGAGGCACTCTCACTGCTTCCAGACATTTTAC 1920  
Db 2201 AGCCTGTGTACTTCTCTCCAATGAGGCACTCTCACTGCTTCCAGACATTTTAC 2260  
Qy 1921 CAAAGTCTGCATGAACACAGTAGGAAAGTTATTCAGTATGGCATCTTACAGTGGCAGAG 1980  
Db 2261 CAAAGTCTGCATGAACACAGTAGGAAAGTTATTCAGTATGGCATCTTACAGTGGCAGAG 2320  
Qy 1981 CACGATGACCAAGGAATATCAGTCTTGTGTCAGCAGCAGTGGGACCAAGAGCTT 2040  
Db 2321 CACGATGACCAAGGAATATCAGTCTTGTGTCAGCAGCAGTGGGACCAAGAGCTT 2380  
Qy 2041 CCAGAACCTTTGCTTGGGAAGTGAAGAGATGAAGACAGTCACTTTTGGGAGGAA 2100  
Db 2381 CCAGAACCTTTGCTTGGGAAGTGAAGAGATGAAGACAGTCACTTTTGGGAGGAA 2440  
Qy 2101 CAGCGAGATGCTTACTCTGAAGTGAAGCAATCCAGGAGCACAAGAGTGTATCACCTTC 2160  
Db 2441 CAGCGAGATGCTTACTCTGAAGTGAAGCAATCCAGGAGCACAAGAGTGTATCACCTTC 2500  
Qy 2161 TTACAGAGACTCCTTTGGGCTTTGTGAGGCTTACAGCTCTGCTGCCATCTTTGTTCAC 2220  
Db 2501 TTACAGAGACTCCTTTGGGCTTTGTGAGGCTTACAGCTCTGCTGCCATCTTTGTTCAC 2560





XX WPI; 2003-229583/22.  
DR P-PSDB; AAO27073.  
XX  
XX Human mitochondrial sn-glycerol-3-phosphate acyltransferase and  
PT antagonists for treatment and prevention of diabetic complications  
XX  
XX Claim 4; Page 33-40; 56pp; Japanese.  
XX  
XX The invention relates to a novel protein having human mitochondrial sn-  
CC glycerol-3-phosphate acyltransferase (GPAM) activity. The novel protein  
CC with GPAM activity can be used in the prevention and treatment of  
CC diabetic complications, including retinopathy and neuropathy, by  
CC administration of antagonists to human GPAM. This polynucleotide sequence  
CC represents a DNA sequence encoding a protein relating to the GPAM  
XX activity protein of the invention.  
XX  
SQ Sequence 2487 BP; 704 A; 549 C; 571 G; 663 T; 0 other;

Query Match 99.9%; Score 2480.8; DB 25; Length 2487;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2482; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGATGAATCTGCACTGACCTTGCTGATACATAGATGTTCTTATCTGCCACATTCATCA 60  
DB 1 ATGGATGAATCTGCACTGACCTTGCTGATACATAGATGTTCTTATCTGCCACATTCATCA 60  
QY 61 GAATACAGTGTGGTGCAGTGAAGCACACAAAGTGAAGGAATGGGTGAGTGGCTTTAGA 120  
DB 61 GAATACAGTGTGGTGCAGTGAAGCACACAAAGTGAAGGAATGGGTGAGTGGCTTTAGA 120  
QY 121 CCCACCGCTTTCAGATCTGCACTTTAAATGGAAGAAAGCCTTAATCAGTCGGAAGAGG 180  
DB 121 CCCACCGCTTTCAGATCTGCACTTTAAATGGAAGAAAGCCTTAATCAGTCGGAAGAGG 180  
QY 181 CCATTTGTTGAAGATGTTTACTCTCGACTCCCGAGAGTGGGACAAATTTTCAAC 240  
DB 181 CCATTTGTTGAAGATGTTTACTCTCGACTCCCGAGAGTGGGACAAATTTTCAAC 240  
QY 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATCAATGAACCTCACACAAGA 300  
DB 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATCAATGAACCTCACACAAGA 300  
QY 301 CACCGGATGGCTTTCGAAGAGCGCTTCTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTG 360  
DB 301 CACCGGATGGCTTTCGAAGAGCGCTTCTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTG 360  
QY 361 CATAGGCGATGTTTGGCCCAACATGTGATGATGAAATGTGCTGAACAGCAGTAGAGTACAA 420  
DB 361 CATAGGCGATGTTTGGCCCAACATGTGATGATGAAATGTGCTGAACAGCAGTAGAGTACAA 420  
QY 421 GAGGCAATTGCAAGATGGCTGCTGAATTAACCTGATGTTCTGCCCAGCAGCAATCA 480  
DB 421 GAGGCAATTGCAAGATGGCTGCTGAATTAACCTGATGTTCTGCCCAGCAGCAATCA 480  
QY 481 AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTTTCAAGAAATGGTTGCCACT 540  
DB 481 AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTTTCAAGAAATGGTTGCCACT 540  
QY 541 GTCTCACCGCAATGATCAGATGCTGCTGGGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
DB 541 GTCTCACCGCAATGATCAGATGCTGCTGGGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 601 TTTTGGACATTCAAATTCAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACG 660  
DB 601 TTTTGGACATTCAAATTCAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACG 660  
QY 661 AATTGCGCTTCTGTTTCTACCACTTCATAGATCCCATATTTGATCTGCTGCTGCTGCTGCTGCT 720  
DB 661 AATTGCGCTTCTGTTTCTACCACTTCATAGATCCCATATTTGATCTGCTGCTGCTGCTGCTGCT 720  
QY 721 TTCAATTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
DB 721 TTCAATTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

DB 721 TTCAATTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
QY 781 ATCCCAATCTTCAGTACCTTGATCATAAGCTTGGGGCTTCTTTCATACGACGAGGCTC 840  
DB 781 ATCCCAATCTTCAGTACCTTGATCATAAGCTTGGGGCTTCTTTCATACGACGAGGCTC 840  
QY 841 GATGAAACACACAGATGGAACGAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
DB 841 GATGAAACACACAGATGGAACGAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
QY 901 GTTGAATTTACTTCCACAGCAGCAATTTCTGGAGATCTTCTGGAGGACACAGCTTCTAGG 960  
DB 901 GTTGAATTTACTTCCACAGCAGCAATTTCTGGAGATCTTCTGGAGGACACAGCTTCTAGG 960  
QY 961 AGTGGAAACCTCTTGTGCTCGGCGAGGACTTTTGTGTCAGTTGTGTAGATCTGCTGCT 1020  
DB 961 AGTGGAAACCTCTTGTGCTCGGCGAGGACTTTTGTGTCAGTTGTGTAGATCTGCTGCT 1020  
QY 1021 ACCAATGTCTATCCAGACATCTTGATTAATACCTGTTTGGAAATCTCTATGATCGCATATC 1080  
DB 1021 ACCAATGTCTATCCAGACATCTTGATTAATACCTGTTTGGAAATCTCTATGATCGCATATC 1080  
QY 1081 GAAGGTCACTACAATGTTGAACAACTGGGCAACCTTAAGAAATGAGAGCTGTGGAGT 1140  
DB 1081 GAAGGTCACTACAATGTTGAACAACTGGGCAACCTTAAGAAATGAGAGCTGTGGAGT 1140  
QY 1141 GTAGCAAGAGGTGTTATTAGAACTGTACGAAACAACTATGTTGTTGTTCCGAGTGGATTTT 1200  
DB 1141 GTAGCAAGAGGTGTTACTAGAAATGTTAGAAACAACTATGTTGTTGTTCCGAGTGGATTTT 1200  
QY 1201 GCACAGCATTTTCTTAAAGGAAATTTTAAAGAGCCAAAGTCAAGAACCCGGTGTCTGCT 1260  
DB 1201 GCACAGCATTTTCTTAAAGGAAATTTTAAAGAGCCAAAGTCAAGAACCCGGTGTCTGCT 1260  
QY 1261 CTACTTTCCCTGGAGAGCGTGTGTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320  
DB 1261 CTACTTTCCCTGGAGAGCGTGTGTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320  
QY 1321 GCTGATCAAGGTGAGACACGTCCTCAATTAAGTGTCCAGAAATGCAACAGATGAATCCCTA 1380  
DB 1321 GCTGATCAAGGTGAGACACGTCCTCAATTAAGTGTCCAGAAATGCAACAGATGAATCCCTA 1380  
QY 1381 CGAAGAGGTTGATTGCAAAATCTGGCTGAGCATATTTCTATTTCACTGTAGCAAGTCTGCT 1440  
DB 1381 CGAAGAGGTTGATTGCAAAATCTGGCTGAGCATATTTCTATTTCACTGTAGCAAGTCTGCT 1440  
QY 1441 GCCATTATGTCACACACATTTGCGCTTGGCTGCTCTCTACACACACAGCAGGGAAT 1500  
DB 1441 GCCATTATGTCACACACATTTGCGCTTGGCTGCTCTCTCTACACACACAGCAGGGAAT 1500  
QY 1501 GATCTCTCCATTTGGTCGAGAGCTTCTTGTGATGAAGAGGAGTCCCTGGCTCGTAT 1560  
DB 1501 GATCTCTCCATTTGGTCGAGAGCTTCTTGTGATGAAGAGGAGTCCCTGGCTCGTAT 1560  
QY 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAGATCTAGTAAATGATGTCATGCCATACAGCTGCTG 1620  
DB 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAGATCTAGTAAATGATGTCATGCCATACAGCTGCTG 1620  
QY 1621 GGAATTTGTGTCAACATCAACCCACCTAGCAGGAAACGATGATTTTTTATCACCCCAAGC 1680  
DB 1621 GGAATTTGTGTCAACATCAACCCACCTAGCAGGAAACGATGATTTTTTATCACCCCAAGC 1680  
QY 1681 ACACCTGCTCCATTCAGTCTTTCGAACTCAACTTCTACAGCAATGGGGTACTTCTATGCTCTT 1740  
DB 1681 ACACCTGCTCCATTCAGTCTTTCGAACTCAACTTCTACAGCAATGGGGTACTTCTATGCTCTT 1740  
QY 1741 ATCATGAGGCGCATCATAGCTTTCAGCTTTTATGAGTCTTCTGAAACAGAGGGGACTGGGG 1800  
DB 1741 ATCATGAGGCGCATCATAGCTTTCAGCTTTTATGAGTCTTCTGAAACAGAGGGGACTGGGG 1800  
QY 1801 GGTCCCACTAGCACCCCACTTAACCTGATCAGCAGGAGCAGCTGGTCGGAAGCGCGCC 1860  
DB 1801 GGTCCCACTAGCACCCCACTTAACCTGATCAGCAGGAGCAGCTGGTCGGAAGCGCGCC 1860



```
QY 1861 AGCTGTGCTACCTTCTCTCAATGAAGGACCACTCTCACTGCTTGGCCAGACATTTTAC 1920
Db |||||
QY 1861 AGCTGTGCTACCTTCTCTCAATGAAGGACCACTCTCACTGCTTGGCCAGACATTTTAC 1920
Db |||||
QY 1921 CAAGTCTGCCATGAACAGTAGGAAGTTTATCAGATATGGCATTTTACAGTGGCAGAG 1980
Db |||||
QY 1921 CAAGTCTGCCATGAACAGTAGGAAGTTTATCAGATATGGCATTTTACAGTGGCAGAG 1980
QY 1981 CAGGATGACAGGAGATATCAGTCTCTAGTCTTCTGAGCAGCAGTGGGACAGAGCTT 2040
Db |||||
QY 1981 CAGGATGACAGGAGATATCAGTCTCTAGTCTTCTGAGCAGCAGTGGGACAGAGCTT 2040
QY 2041 CCAGAACCTTTGCTTCTGGAGAAGTATGAAGAAGATGAACAGCAGTGAATTTGGGGAGGAA 2100
Db |||||
QY 2041 CCTGAACCTTTGCTTCTGGAGAAGTATGAAGAAGATGAACAGCAGTGAATTTGGGGAGGAA 2100
QY 2101 CAGGAGATTTGCTACCTGAAGTGGAGCCAAATCAAGGAGCACCAGCAGTTTATCACTTC 2160
Db |||||
QY 2101 CAGGAGATTTGCTACCTGAAGTGGAGCCAAATCAAGGAGCACCAGCAGTTTATCACTTC 2160
QY 2161 TTACAGAGCTCTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTAC 2220
Db |||||
QY 2161 TTACAGAGCTCTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTAC 2220
QY 2221 AACTTCAGTGTCTGCTTCTCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTAATA 2280
Db |||||
QY 2221 AACTTCAGTGTCTGCTTCTCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTAATA 2280
QY 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCACATATTTCTTGTGAAG 2340
Db |||||
QY 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCACATATTTCTTGTGAAG 2340
QY 2341 AATGCTGTGAATATGTTAAGGATATTTGGGTTTTCAGGAGACCAACAAAGAGAGTG 2400
Db |||||
QY 2341 AATGCTGTGAATATGTTAAGGATATTTGGGTTTTCAGGAGACCAACAAAGAGAGTG 2400
QY 2401 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACGACAAACAACTTCTAGAA 2460
Db |||||
QY 2401 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACGACAAACAACTTCTAGAA 2460
QY 2461 TATATTCTGAGTTTGTGGTGCTG 2484
Db |||||
QY 2461 TATATTCTGAGTTTGTGGTGCTG 2484

RESULT 5
AAD35221
AC AAD35221 standard; cDNA; 3210 BP.
AC AAD35221;
DE 25-JUL-2002 (first entry)
XX Human TRNFR-6 cDNA.
XX Human; transferase; developmental disorder; cell proliferative disorder;
XX TRNFR-6; neurological disorder; autoimmune disorder; parasitic infection;
XX inflammatory disorder; endocrine; antiparasitic; immunosuppressive;
XX cytostatic; neurological; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 338..2824
XX FT /*tag= a
XX FT /product= "Human TRNFR-6 protein"
XX PN WO200226950-A2.
XX PD 04-APR-2002.
XX PF 28-SEP-2001; 2001WO-US30424.
```

```
XX 29-SEP-2000; 2000US-236523P.
PR 06-OCT-2000; 2000US-238481P.
PR 27-OCT-2000; 2000US-244025P.
PR 03-NOV-2000; 2000US-246001P.
PR 09-NOV-2000; 2000US-247931P.
PR 16-NOV-2000; 2000US-249639P.
PR 21-NOV-2000; 2000US-252819P.
XX (INCY-) INCYTE GENOMICS INC.
XX Lal PG, Tang YT, Yue H, Burford N, Gandhi AR, Warren BA, Yao MG;
PI Tribouley CM, Baughn MR, Lee EA, Hafalia AJA, Lu Y, Griffin JA;
PI Sanjanwala MS, Ding L;
XX WPI: 2002-362492/39.
DR P-PSDB; AAE22144.
XX Novel human transferase polypeptides and polynucleotides, useful in
PT treating e.g., cell proliferative and autoimmune disorders -
XX Claim 81; Page 154-155; 168pp; English.
XX The present invention relates to novel human transferases (TRNFR) and
CC polynucleotides encoding such proteins. The TRNFR proteins are useful
CC for treating disorders associated with a decreased expression of
CC functional TRNFR, e.g., cell proliferative, developmental, neurological,
CC autoimmune/inflammatory disorders and parasitic infections. Antagonists
CC of TRNFR proteins are useful in treating disorders associated with
CC increased activity of TRNFR. The present sequence is human TRNFR-6 cDNA.
XX SQ Sequence 3210 BP; 896 A; 721 C; 735 G; 858 T; 0 other;
```

```
Query Match 99.9%; Score 2480.8; DB 24; Length 3210;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2482; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGATGAATCTGCACTGACCTTGGTCAATAGATGTTTCTTATCTGCCACATTCATCA 60
Db 338 ATGGATGAATCTGCACTGACCTTGGTCAATAGATGTTTCTTATCTGCCACATTCATCA 397
QY 61 GAATACAGTGTGGTGCATGTAAGCACACAGTGGAGGAATGGGTGAGTGTGGCTTTAGA 120
Db 398 GAATACAGTGTGGTGCATGTAAGCACACAGTGGAGGAATGGGTGAGTGTGGCTTTAGA 457
QY 121 CCCACGCTTTCAGATCTGCAACTTTAAATCGAAGAAAGCCTAATGAGTCGGAAGG 180
Db 458 CCCACCATCTTCAGATCTGCAACTTTAAATCGAAGAAAGCCTAATGAGTCGGAAGG 517
QY 181 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCAGAGCTGGGACAAATTTTCAAC 240
Db 518 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCAGAGCTGGGACAAATTTTCAAC 577
QY 241 CCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAACTCACACAAG 300
Db 578 CCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAACTCACACAAG 637
QY 301 CACCGGGATGCTTCAGAGCCCTTTCTTAGCTTCTTTTATTCAAGAGCGAGATGTG 360
Db 638 CACCGGGATGCTTCAGAGCCCTTTCTTAGCTTCTTTTATTCAAGAGCGAGATGTG 697
QY 361 CATAGGGCATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 420
Db 698 CATAGGGCATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 757
QY 421 GAGGCAATTCAGAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCAGCAGCAATCA 480
Db 758 GAGGCAATTCAGAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCAGCAGCAATCA 817
QY 481 AAGCCGTTTAAACAAAGTGAAGAAAGCTAAAGGATTTCTTCAGAAATGGTTGCCACT 540
Db 818 AAGCCGTTTAAACAAAGTGAAGAAAGCTAAAGGATTTCTTCAGAAATGGTTGCCACT 877
```



QY 541 GTCTCACCGGCAATGATCAGACTGACTGGTGGGTGCTGCTAAACTGTTTCAACAGCTTC 600  
Db 878 GTCTCACCGGCAATGATCAGACTGACTGGTGGGTGCTGCTAAACTGTTTCAACAGCTTC 937  
QY 601 TTTTGGAAATCAAAATTCACAAAGGTCAACTTTGAGATGGTTAAAGCTGCAACTGAGACG 660  
Db 938 TTTTGGAAATCAAAATTCACAAAGGTCAACTTTGAGATGGTTAAAGCTGCAACTGAGACG 997  
QY 661 AATTTCGCGCTTCTGTTTCTACAGTTCATAGATCCCATATTTGACTATCTGCTGCTCACT 720  
Db 998 AATTTCGCGCTTCTGTTTCTACAGTTCATAGATCCCATATTTGACTATCTGCTGCTCACT 1057  
QY 721 TTCATTCTCTTCTGTCATTAACATCAAGACCATACATTTGCTCAGGCAATTAATCTCAAC 780  
Db 1058 TTCATTCTCTTCTGTCATTAACATCAAGACCATACATTTGCTCAGGCAATTAATCTCAAC 1117  
QY 781 ATCCCAATCTTCAGTACCTTGATCCATAAGCTTTGGGGCTTCTTCATACGACGAGGCTC 840  
Db 1118 ATCCCAATCTTCAGTACCTTGATCCATAAGCTTTGGGGCTTCTTCATACGACGAGGCTC 1177  
QY 841 GATGAACAACCCAGATGGAACGAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db 1178 GATGAACAACCCAGATGGAACGAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1237  
QY 901 GTTGAATTTACTCGACAGCAGCAATTTCTTGAGATCTTCTGGAAGGCACAGTTCTAGG 960  
Db 1238 GTTGAATTTACTCGACAGCAGCAATTTCTTGAGATCTTCTGGAAGGCACAGTTCTAGG 1297  
QY 961 AGTGGAAAACTCTTTGTGCTCGGCAGGACTTTTGTGAGTTGTTGTTGAGTACTCTGTCT 1020  
Db 1298 AGTGGAAAACTCTTTGTGCTCGGCAGGACTTTTGTGAGTTGTTGTTGAGTACTCTGTCT 1357  
QY 1021 ACCAATGTCATCCACAGACATCTTGATAATACCTGTTGGAATCTCTATGATCGCATATC 1080  
Db 1358 ACCAATGTCATCCACAGACATCTTGATAATACCTGTTGGAATCTCTATGATCGCATATC 1417  
QY 1081 GAAGTCTACTCAATGTTGAACAACTGGGCAAACTTAAGAAAGATGAGAGCTGTGGAGT 1140  
Db 1418 GAAGTCTACTCAATGTTGAACAACTGGGCAAACTTAAGAAAGATGAGAGCTGTGGAGT 1477  
QY 1141 GTACGAGAGTGTATTAGAAATGTTAGAAATACTATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200  
Db 1478 GTAGCAAGAGTGTATTAGAAATGTTAGAAATACTATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1537  
QY 1201 GCACAGCCATTTCTTAAAGGAATATTAGAAAGCCAAAGTTCAGAAACCCGTTCTGCT 1260  
Db 1538 GCACAGCCATTTCTTAAAGGAATATTAGAAAGCCAAAGTTCAGAAACCCGTTCTGCT 1597  
QY 1261 CTACTTTCCCTGGAGCAAGCGTTGTTACAGCTATACCTTCCCTTCAAGACCCAGTGTGCT 1320  
Db 1598 CTACTTTCCCTGGAGCAAGCGTTGTTACAGCTATACCTTCCCTTCAAGACCCAGTGTGCT 1657  
QY 1321 GCTGATGAAGTGAAGACAGCTCCATTAATGAGTCCAGAAATGCAAGATGAATCCCTA 1380  
Db 1658 GCTGATGAAGTGAAGACAGCTCCATTAATGAGTCCAGAAATGCAAGATGAATCCCTA 1717  
QY 1381 CGAAGAGGTTGATTGCAAAATCTGCTGAGCATATTCTATTCACTGCTAGCAAGTCTGCT 1440  
Db 1718 CGAAGAGGTTGATTGCAAAATCTGCTGAGCATATTCTATTCACTGCTAGCAAGTCTGCT 1777  
QY 1441 GCCATTATGTCACACACATTTGGCTTGGCTTCCCTCTACAGACACAGCAGGGAATT 1500  
Db 1778 GCCATTATGTCACACACATTTGGCTTGGCTTCCCTCTACAGACACAGCAGGGAATT 1837  
QY 1501 GATCTCTCAATTTGGTGAAGACTTTCTTTGTGATGAAAGGAAGTCTCGCTCGTGAT 1560  
Db 1838 GATCTCTCAATTTGGTGAAGACTTTCTTTGTGATGAAAGGAAGTCTCGCTCGTGAT 1897  
QY 1561 TTTGACCTGGGTTCTCAGGAATTTCAAGATGTAGTAATGATGCTGCATACAGCTGCTG 1620  
Db 1898 TTTGACCTGGGTTCTCAGGAATTTCAAGATGTAGTAATGATGCTGCATACAGCTGCTG 1957  
QY 1621 GGAAATTTGTGTCACAATCACCCACACTAGCAGGAACGATGAGTTTTTTTATCACCCCCAGC 1680

Db 1958 GGAAATTTGTGTCACAATCACCCACACTAGCAGGAACGATGAGTTTTTTTATCACCCCCAGC 2017  
QY 1681 ACAAATGTCCTCAGTCTTCCGAACTCAACTTCTACAGCAATGGGGTACTTCTATGTCCTT 1740  
Db 2018 ACAAATGTCCTCAGTCTTCCGAACTCAACTTCTACAGCAATGGGGTACTTCTATGTCCTT 2077  
QY 1741 ATCAATGAGGCGCATCATAGCTTGCAGCTTTATGCAGTCTTCTGAACAAGAGGGGACTGGGG 1800  
Db 2078 ATCAATGAGGCGCATCATAGCTTGCAGCTTTATGCAGTCTTCTGAACAAGAGGGGACTGGGG 2137  
QY 1801 GGTCCCACTAGCACCCACCTAATCCTGATCAGCAGAGGAGCTGGTGGGAAGGCGGCC 1860  
Db 2138 GGTCCCACTAGCACCCACCTAATCCTGATCAGCAGGAGGAGCTGGTGGGAAGGCGGCC 2197  
QY 1861 AGCCTGTGCTACCTTCTCTCAATGAAGGACATCTCAGTCCCTTGCACACATTTTAC 1920  
Db 2198 AGCCTGTGCTACCTTCTCTCAATGAAGGACATCTCAGTCCCTTGCACACATTTTAC 2257  
QY 1921 CAACTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCAATTTTACAGTGGCAGAG 1980  
Db 2258 CAACTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCAATTTTACAGTGGCAGAG 2317  
QY 1981 CACGATCAGCAGGAAGATATCAGTCTAGTCTTCTGCTGAGCAGCTGGGAACAAGGCTT 2040  
Db 2318 CACGATCAGCAGGAAGATATCAGTCTAGTCTTCTGCTGAGCAGCTGGGAACAAGGCTT 2377  
QY 2041 CCAGAACTTTGCTCTGAGAAAGTATGAAGATGAAGACAGTGAATTTGGGGAGGAA 2100  
Db 2378 CCTGAACCTTTGCTCTGAGAAAGTATGAAGATGAAGACAGTGAATTTGGGGAGGAA 2437  
QY 2101 CAGCAGATGCTACCTGAAGGTGAGCAATCCAAAGAGCAGCAGCTTTTATCACCTTC 2160  
Db 2438 CAGCAGATGCTACCTGAAGGTGAGCAATCCAAAGAGCAGCAGCTTTTATCACCTTC 2497  
QY 2161 TTACAGAGACTCTTTGGGCTTTGCTGAGGCTTACAGCTCTGCTGCCATCTTTGTTTCCAC 2220  
Db 2498 TTACAGAGACTCTTTGGGCTTTGCTGAGGCTTACAGCTCTGCTGCCATCTTTGTTTCCAC 2557  
QY 2221 AACTTCAGTGTCTGTTCCAGAACCTTGAGTATCTGCAAAAGTTGCACAAATACCTAATA 2280  
Db 2558 AACTTCAGTGTCTGTTCCAGAACCTTGAGTATCTGCAAAAGTTGCACAAATACCTAATA 2617  
QY 2281 ACCAGAACAGAAAGAAATGTTGCAAGTATGCTCAGAGTGCACATATGTTCTTGTGAAG 2340  
Db 2618 ACCAGAACAGAAAGAAATGTTGCAAGTATGCTCAGAGTGCACATATGTTCTTGTGAAG 2677  
QY 2341 AATGCTGTGAAATGTTTAAAGGATATTGGGTTTCAAGGAGACCAAAACAAAGAGAGTG 2400  
Db 2678 AATGCTGTGAAATGTTTAAAGGATATTGGGTTTCAAGGAGACCAAAACAAAGAGAGTG 2737  
QY 2401 TCTGTTTGTAGAACGTGACGACGACTTTTCTACCTCAATGCAACCCGACAAAACCTTCTAGAA 2460  
Db 2738 TCTGTTTGTAGAACGTGACGACGACTTTTCTACCTCAATGCAACCCGACAAAACCTTCTAGAA 2797  
QY 2461 TATATCTGAGTTTGTGCTGCTG 2484  
Db 2798 TATATCTGAGTTTGTGCTGCTG 2821

## RESULT 6

ABX72233

ID ABX72233 standard; cDNA; 2684 BP.

XX

AC ABX72233;

XX

DT 03-JUN-2003 (first entry)

XX

DE Human NOVX polynucleotide #64.

XX

KW Human; NOVX; gene; ss; metabolic disorder; cardiomyopathy; diabetes; ASD;

KW hypertension; congenital heart defect; aortic stenosis; valve disease;

KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;

KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;  
 KW tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;  
 KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;  
 KW Parkinson's disease; immune disorder; haematopoietic disorder;  
 KW haemophilia; hypercoagulation; Crohn's disease; cancer.

XX Homo sapiens.

OS WO200281498-A2.

PN 17-OCT-2002.

PD 03-APR-2002; 2002WO-US10780.

PF 03-APR-2001; 2001US-281086P.

PR 03-APR-2001; 2001US-281136P.

PR 05-APR-2001; 2001US-281863P.

PR 06-APR-2001; 2001US-281906P.

PR 10-APR-2001; 2001US-282020P.

PR 10-APR-2001; 2001US-282930P.

PR 10-APR-2001; 2001US-282934P.

PR 12-APR-2001; 2001US-283512P.

PR 13-APR-2001; 2001US-283710P.

PR 17-APR-2001; 2001US-284234P.

PR 19-APR-2001; 2001US-285325P.

PR 20-APR-2001; 2001US-285381P.

PR 20-APR-2001; 2001US-285609P.

PR 23-APR-2001; 2001US-285748P.

PR 23-APR-2001; 2001US-285890P.

PR 24-APR-2001; 2001US-286068P.

PR 25-APR-2001; 2001US-286292P.

PR 27-APR-2001; 2001US-287213P.

PR 02-MAY-2001; 2001US-288257P.

PR 23-MAY-2001; 2001US-294164P.

PR 30-MAY-2001; 2001US-294484P.

PR 18-JUN-2001; 2001US-298952P.

PR 19-JUN-2001; 2001US-299237P.

PR 19-JUN-2001; 2001US-299276P.

PR 12-SEP-2001; 2001US-318750P.

PR 25-SEP-2001; 2001US-324800P.

PR 25-SEP-2001; 2001US-324802P.

PR 27-SEP-2001; 2001US-325684P.

PR 14-OCT-2001; 2001US-330143P.

PR 14-NOV-2001; 2001US-332131P.

PR 14-NOV-2001; 2001US-332240P.

PR 21-NOV-2001; 2001US-332179P.

PR 21-NOV-2001; 2001US-332115P.

PR 04-DEC-2001; 2001US-337621P.

PR 03-JAN-2002; 2002US-345783P.  
 PR 16-JAN-2002; 2002US-350251P.  
 PR 02-APR-2002; 2002US-0114270.

XX (CURA-) CURAGEN CORP.

XX Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA, Patturajan M;

XX Liu X, Gusev VY, Li L, Vernet CAM, Zerhusen BD, Gorman L;

XX Sheny SG, Pena CE, Smithson G, Burges CE, Gerlach V;

XX Padigar M, Shimkets RA, Gangoli EA, Taupier RJ, Casman SJ, Ji W;

XX Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;

XX Macdougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;

XX Ellerman K;

XX WPI; 2003-046858/04.

DR P-PSDB; ABUS4605.

XX New isolated NOVX polypeptide useful for treating atherosclerosis,

XX metabolic disorders, diabetes, obesity, infectious disease, anorexia,

XX neurodegenerative disorders, Alzheimer's disease and cancer

XX Claim 17; Page 231-232; 666pp; English.

PS The invention relates to human polypeptides, termed NOVX, and the

XX polynucleotides encoding them. The polypeptides and polynucleotides are

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC useful for diagnosing disease, and screening for potential therapeutic  
 CC agents. The sequences are useful for treating metabolic disorders,  
 CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic  
 CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,  
 CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular  
 CC septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,  
 CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative  
 CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,  
 CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease  
 CC and cancer. Sequences ABX72170-ABX72275 represent human NOVX  
 CC polynucleotides of the invention.

XX

SQ Sequence 2684 BP; 754 A; 583 C; 615 G; 732 T; 0 other;

Query Match 99.7%; Score 2476; DB 25; Length 2684;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2479; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGATGAATCTGCACTGACCTTGGTACAATAGATGTTTCTTATCTGCCACATTCATCA 60

DB 101 ATGGATGAATCTGCACTGACCTTGGTACAATAGATGTTTCTTATCTGCCACATTCATCA 160

QY 61 GAATACAGTGTGGTGCATGTAAGCACACAAGTAGGAATGGGTGAGTGTGGCTTTAGA 120

DB 161 GAATACAGTGTGGTGCATGTAAGCACACAAGTAGGAATGGGTGAGTGTGGCTTTAGA 220

QY 121 CCCACCGTCTCAGATCTGCACTTAAATGGAAGAAAGCTAATAGTCGGAAGG 180

DB 221 CCCACCATCTTCAGATCTGCACTTAAATGGAAGAAAGCTAATAGTCGGAAGG 280

QY 181 CCATTTGTTGGAAGATGTTTACTCTGCACTCCCAGAGCTGGGACAAAATTTTCAAC 240

DB 281 CCATTTGTTGGAAGATGTTTACTCTGCACTCCCAGAGCTGGGACAAAATTTTCAAC 340

QY 241 CCCAGTATCCCTCTTGGGTTTGGGAATGTTATATCAATGAATCACTACAAGA 300

DB 341 CCCAGTATCCCTCTTGGGTTTGGGAATGTTATATCAATGAATCACTACAAGA 400

QY 301 CACCGCGATGCTTGCAGAGCGCTTCTTCTAGTCTTTTATCAAGCGAGATGTG 360

DB 401 CACCGCGATGCTTGCAGAGCGCTTCTTCTAGTCTTTTATCAAGCGAGATGTG 460

QY 361 CATAAGGCGATGTTTGCCACCAATGCTGCTGAAAATGCTGCAACAGCAGTAGATACAA 420

DB 461 CATAAGGCGATGTTTGCCACCAATGCTGCTGAAAATGCTGCAACAGCAGTAGATACAA 520

QY 421 GAGGCAATTCAGAAAGTGGCTGCTGAATTAACCTGTAGTGGTCTGCCAGCAGCAATCA 480

DB 521 GAGGCAATTCAGAAAGTGGCTGCTGAATTAACCTGTAGTGGTCTGCCAGCAGCAATCA 580

QY 481 AAAGCGTTTAAACAAGTGAAAAGAAAGCTAAAGGATTTCTCAAGAAATGGTGGCACT 540

DB 581 AAAGCGTTTAAACAAGTGAAAAGAAAGCTAAAGGATTTCTCAAGAAATGGTGGCACT 640

QY 541 GTCTCACCGCGATGATCAGATGCTGCTGGTGGGTGCTGCTAAACCTGTTCAACAGCTTC 600

DB 641 GTCTCACCGCGATGATCAGATGCTGCTGGTGGGTGCTGCTAAACCTGTTCAACAGCTTC 700

QY 601 TTTTGGAAATCAAAATTCACAAAGTCAACTGAGATGGTTAAAGCTGCAACTGAGACG 660

DB 701 TTTTGGAAATCAAAATTCACAAAGTCAACTGAGATGGTTAAAGCTGCAACTGAGACG 760

QY 661 AATTTGCCGCTTCTGTTTCTTACAGTTCATAGTCCATATTCAGTATCTGCTGCTCACT 720

DB 761 AATTTGCCGCTTCTGTTTCTTACAGTTCATAGTCCATATTCAGTATCTGCTGCTCACT 820

QY 721 TTTTATTTCTTCTGCTGCTTAAACCAACCAATCAATGCTTCAAGCAATATCTCAAC 780

DB 821 TTTTATTTCTTCTGCTGCTTAAACCAACCAATCAATGCTTCAAGCAATATCTCAAC 880

QY 781 ATCCCAATCTTCAAGTACCTTGTATCCATAGCTTGGGGGCTTCTTATACGACGAAGCTC 840

DB 881 ATCCCAATCTTCAAGTACCTTGTATCCATAGCTTGGGGGCTTCTTATACGACGAAGCTC 940

QY 841 GATGAACACACGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db |||||  
QY 941 GATGAACACACGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1000  
Db |||||  
QY 901 GTTGAATTTACTTCGACACGACCAATTTCTTGAGATCTTCTGGAGGACACAGTCTTAGG 960  
Db |||||  
QY 1001 GTTGAATTTACTTCGACACGACCAATTTCTTGAGATCTTCTGGAGGACACAGTCTTAGG 1060  
Db |||||  
QY 961 AGTGAATAAACCCTTTGTGCTCGGACGAGCTTTTGTGCTAGTGTGGTAGATCTCTGTCT 1020  
Db |||||  
QY 1061 AGTGAATAAACCCTTTGTGCTCGGACGAGCTTTTGTGCTAGTGTGGTAGATCTCTGTCT 1120  
Db |||||  
QY 1021 ACCAATGTCATCCAGACATCTTGATATACTGTTTGAATCTCTTATGATCGAATATC 1080  
Db |||||  
QY 1121 ACCAATGTCATCCAGACATCTTGATATACTGTTTGAATCTCTTATGATCGAATATC 1180  
Db |||||  
QY 1081 GAAGTCACTCAATGTTGAACACTGGGCAACCTTAAGAAATGAGAGCTGTGGAGT 1140  
Db |||||  
QY 1181 GAAGTCACTCAATGTTGAACACTGGGCAACCTTAAGAAATGAGAGCTGTGGAGT 1240  
Db |||||  
QY 1141 GTAGCAAGAGTGTATTAGAAATGTTACGAAAAAATATGTTGTCTCCGAGTGGATTTT 1200  
Db |||||  
QY 1241 GTAGCAAGAGTGTATTAGAAATGTTACGAAAAAATATGTTGTCTCCGAGTGGATTTT 1300  
Db |||||  
QY 1201 GCACAGCCATTTCTTTAAAGAAATATTGAAAGCCAAAGTTCAGAAACCCGGTGTCTGCT 1260  
Db |||||  
QY 1301 GCACAGCCATTTCTTTAAAGAAATATTGAAAGCCAAAGTTCAGAAACCCGGTGTCTGCT 1360  
Db |||||  
QY 1261 CTACTTTCCCTGGAGCAGCTGTTTACGAGTATCTTCTTCAAGACCCAGTGTGCT 1320  
Db |||||  
QY 1361 CTACTTTCCCTGGAGCAGCTGTTTACGAGTATCTTCTTCAAGACCCAGTGTGCT 1420  
Db |||||  
QY 1321 GCTGATGAAGTGTAGACACAGCTTCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db |||||  
QY 1421 GCTGATGAAGTGTAGACACAGCTTCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1480  
Db |||||  
QY 1381 CGAAGGAGTGTGTTGCAAAATCTGGCTGAGCATATTTCTATTTCACTGTAGCAAGTCTCTGT 1440  
Db |||||  
QY 1481 CGAAGGAGTGTGTTGCAAAATCTGGCTGAGCATATTTCTATTTCACTGTAGCAAGTCTCTGT 1540  
Db |||||  
QY 1441 GCCATTATGTCACACACATTTGGCTTGGCTGCTCTCTACACACACAGCAGCGGAATT 1500  
Db |||||  
QY 1541 GCCATTATGTCACACACATTTGGCTTGGCTGCTCTCTACACACACAGCAGCGGAATT 1600  
Db |||||  
QY 1501 GATCTCTCCATTTGGTCTGAAGACTTCTTTGTGATGAAGAGGAAGTCTTGGCTCTGTGAT 1560  
Db |||||  
QY 1601 GATCTCTCCATTTGGTCTGAAGACTTCTTTGTGATGAAGAGGAAGTCTTGGCTCTGTGAT 1660  
Db |||||  
QY 1561 TTTGACCTGGGGTTCTCAGGAAATTCAGAAGATGTAGTAATGTCATGCCATACAGCTGCTG 1620  
Db |||||  
QY 1661 TTTGACCTGGGGTTCTCAGGAAATTCAGAAGATGTAGTAATGTCATGCCATACAGCTGCTG 1720  
Db |||||  
QY 1621 GGAATTTGTTCACAAATCACCACACTAGCAGGAACGATGATGTTTTTATCACCACCCAGC 1680  
Db |||||  
QY 1721 GGAATTTGTTCACAAATCACCACACTAGCAGGAACGATGATGTTTTTATCACCACCCAGC 1780  
Db |||||  
QY 1681 ACAATGTCCTCAGTCTCGAACTTCTGAACTTCTACAGCAATGGGTACTTCATGCTCTTT 1740  
Db |||||  
QY 1781 ACAATGTCCTCAGTCTCGAACTTCTGAACTTCTACAGCAATGGGTACTTCATGCTCTTT 1840  
Db |||||  
QY 1741 ATCATGGAGGCCATCATAGCTTTCAGGCTTTTATGCAAGTCTTGAACAAGAGGGGACTGGGG 1800  
Db |||||  
QY 1841 ATCATGGAGGCCATCATAGCTTTCAGGCTTTTATGCAAGTCTTGAACAAGAGGGGACTGGGG 1900  
Db |||||  
QY 1801 GGTCCCACTAGCACCCCACTTAACTGATACGAGGAGCAGCTGGTTCGGAAGCGCGCC 1860  
Db |||||  
QY 1901 GGTCCCACTAGCACCCCACTTAACTGATACGAGGAGCAGCTGGTTCGGAAGCGCGCC 1960  
Db |||||  
QY 1861 AGCTGTGCTACCTTCTTCAATGAAGGACCACTCTCACTGCTTTCGACAGATTTTAC 1920  
Db |||||  
QY 1961 AGCTGTGCTACCTTCTTCAATGAAGGACCACTCTCACTGCTTTCGACAGATTTTAC 2020  
Db |||||

RESULT 7  
ABZ12025

ID ABZ12025 standard; cDNA; 3273 BP.

XX AC ABZ12025;

XX AC ABZ12025;

XX DT 20-JAN-2003 (first entry)

XX DE Human polynucleotide SEQ ID NO 907.

Human; genome mapping; gene therapy; food supplement; virus; fungus;  
cell-proliferative disorder; neurodegenerative disease; bacterial;  
Parkinson's disease; Alzheimer's disease; autoimmune disease;  
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;  
antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
hemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
antiarthritic; gene; ss.

XX OS Homo sapiens.

XX PN W0200270539-A2.

XX PN 12-SEP-2002.

XX PF 05-MAR-2002; 2002WO-US05095.

XX PF 05-MAR-2001; 2001US-0799451.

XX PR (HYSE-) HYSEQ INC.

XX PA Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;

XX PI

PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Wehrman T, Wang J, Wang D, Drmanac RT;  
XX WPI: 2002-759812/82.  
DR P-PSDB; ABP69808.

XX New polynucleotides comprising sequences assembled from expressed  
PT sequence tags (ESTs), useful for treating cell-proliferative,  
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or  
PT platelet or coagulation disorders -

XX Claim 1; SEQ ID NO 907; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising a  
CC nucleotide sequence selected from any of 948 sequences  
CC (ABZ11119-ABZ12066) or their mature protein coding portion, active domain  
CC coding protein or complementary sequences. The polynucleotides are useful  
CC for identifying expressed genes or for physical mapping of human genome.  
CC The encoded polypeptides (ABP6902-ABP69849) are useful as molecular  
CC weight markers, as a food supplement, for generating antibodies, in  
CC medical imaging, screening and diagnostic assays and for treating  
CC cell-proliferative disorders (cancer), neurodegenerative diseases  
CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple  
CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid  
CC disorders, platelet or coagulation disorders, wound, burns, incision,  
CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,  
CC parasitic), arthritis, etc.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 3273 BP; 915 A; 704 C; 733 G; 921 T; 0 other;

Query Match 88.0%; Score 2185.8; DB 24; Length 3273;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2187; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 296 CAAGACACCGCGATGGCTTGCAGACGCTTCTTACGTTCTTTTATTCAAGACCGAG 355  
DB 173 CCAGACACCGCGATGGCTTGCAGACGCTTCTTACGTTCTTTTATTCAAGACCGAG 232  
QY 356 ATGTGCAATAGGCGATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAG 415  
DB 233 ATGTGCAATAGGCGATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAG 292  
QY 416 TACAAGAGCAATTCAGAAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCAGCAGC 475  
DB 293 TACAAGAGCAATTCAGAAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCAGCAGC 352  
QY 476 AATCAAAAGCCGTTAACAAGTGAAAGAGCTAAAGAGTCTTCAAGAAATGGTTG 535  
DB 353 AATCAAAAGCCGTTAACAAGTGAAAGAGCTAAAGAGTCTTCAAGAAATGGTTG 412  
QY 536 CCACTGTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGTCTTAAACCTGTTCAACA 595  
DB 413 CCACTGTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGTCTTAAACCTGTTCAACA 472  
QY 596 GCTTCTTTTGGAAACATTCAAAATTCAAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTG 655  
DB 473 GCTTCTTTTGGAAACATTCAAAATTCAAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTG 532  
QY 656 AGACGAATTTGGCGTCTCTGTTTCTACAGTTTCATAGATCCCATATTTGACTATCTGCTGC 715  
DB 533 AGACGAATTTGGCGTCTCTGTTTCTACAGTTTCATAGATCCCATATTTGACTATCTGCTGC 592  
QY 716 TCACCTTTTCAATCTCTCTGCCATAACATCAAAAGCACCACATACATTTGCTTCAGGCAATAATC 775  
DB 593 TCACCTTTTCAATCTCTCTGCCATAACATCAAAAGCACCACATACATTTGCTTCAGGCAATAATC 652  
QY 776 TCACGATCCCAATCTTCAGTACCTTGATCCGATAGCTTGGGGCTCTTCTCATACGACGAA 835  
DB 653 TCACGATCCCAATCTTCAGTACCTTGATCCGATAGCTTGGGGCTCTTCTCATACGACGAA 712

QY 836 GGCTCGATGAACACACAGATGACCGGAAAGATGTTTCTTATAGAGCTTTGCTCCATGGGC 895  
DB 713 GGCTCGATGAACACACAGATGACCGGAAAGATGTTTCTTATAGAGCTTTGCTCCATGGGC 772  
QY 896 ATATAGTTGAATTTACTTTCGACAGAGCAATTTCTTGAGATCTTCTCGAAGGACACAGTT 955  
DB 773 ATATAGTTGAATTTACTTTCGACAGAGCAATTTCTTGAGATCTTCTCGAAGGACACAGTT 832  
QY 956 CTAGAGATGGAAGAAACCTTCTGCTCGGGCAGGACTTTTGTGAGTTGTTGTTAGATCTC 1015  
DB 833 CTAGAGATGGAAGAAACCTTCTGCTCGGGCAGGACTTTTGTGAGTTGTTGTTAGATCTC 892  
QY 1016 TGTCTACCAATGCTATCCAGACATCTTGTATTAATACCTGTTGGAAATCTCTATGATCGCA 1075  
DB 893 TGTCTACCAATGCTATCCAGACATCTTGTATTAATACCTGTTGGAAATCTCTATGATCGCA 952  
QY 1076 TTATCGAAGGTCATCAATGTTGTAACAACTGGGCAACCTAAGAAAGATGAGAGCTGT 1135  
DB 953 TTATCGAAGGTCATCAATGTTGTAACAACTGGGCAACCTAAGAAAGATGAGAGCTGT 1012  
QY 1136 GGAGTGTAGCAAGAGGTGTTATTAGNAATGTTACGAAACAACTATGTTGTTCCGAGTGG 1195  
DB 1013 GGAGTGTAGCAAGAGGTGTTATTAGAAATGTTACGAAACAACTATGTTGTTCCGAGTGG 1072  
QY 1196 ATTTTGCACAGCCATTTTCTTAAAGGAATATTAGAAAGCCAAAGTCAGAAACCGGTGT 1255  
DB 1073 ATTTTGCACAGCCATTTTCTTAAAGGAATATTAGAAAGCCAAAGTCAGAAACCGGTGT 1132  
QY 1256 CTGCTCTACTTTTCTGGAGCAAGGTGTTTACAGCTATATCTTCTTCAAGACCCAGTG 1315  
DB 1133 CTGCTCTACTTTTCTGGAGCAAGGTGTTTACAGCTATATCTTCTTCAAGACCCAGTG 1192  
QY 1316 ATGCTGCTGATGAAGGTAGACACGCTCCATTAATGAGTCCAGAAATCAACAGATGAAT 1375  
DB 1193 ATGCTGCTGATGAAGGTAGACACGCTCCATTAATGAGTCCAGAAATCAACAGATGAAT 1252  
QY 1376 CCCTACGAGGAGGTGATTTGCAAAATCTGGCTGAGCATATTTCTTCTGCTAGCAAGT 1435  
DB 1253 CCCTACGAGGAGGTGATTTGCAAAATCTGGCTGAGCATATTTCTTCTGCTAGCAAGT 1312  
QY 1436 CTGTGCAATTAATGTTCCACACATTTGTTGGCTTGGCTGCTCTCTACAGACACAGCAGG 1495  
DB 1313 CTGTGCAATTAATGTTCCACACATTTGTTGGCTTGGCTGCTCTCTACAGACACAGCAGG 1372  
QY 1496 GAATGATCTCTCCACATTTGTTGGAAGCTTCTTGTGATGAAGAGGAGTCTTGGCTC 1555  
DB 1373 GAATGATCTCTCCACATTTGTTGGAAGCTTCTTGTGATGAAGAGGAGTCTTGGCTC 1432  
QY 1556 GTGATTTTGAACCTGGGTTCTCAGGAAATTCAGAAAGTGTAGTAATGCATGCATACAGC 1615  
DB 1433 GTGATTTTGAACCTGGGTTCTCAGGAAATTCAGAAAGTGTAGTAATGCATGCATACAGC 1492  
QY 1616 TGCTGGGAAATTTGTCAAAATCAACCACTAGCAGGAAACGATGAGTTTTTATATCACCC 1675  
DB 1493 TGCTGGGAAATTTGTCAAAATCAACCACTAGCAGGAAACGATGAGTTTTTATCACCC 1552  
QY 1676 CCAGACACTCTGCCATCAGTCTTGGAACTCACTTCTACAGCAATGGGTACTTCTG 1735  
DB 1553 CCAGACACTCTGCCATCAGTCTTGGAACTCACTTCTACAGCAATGGGTACTTCTG 1612  
QY 1736 TCTTTATCATGAGGCGCATCATAGCTTGCAGGCTTTATGCACTTCTGAAACAGAGGGGAC 1795  
DB 1613 TCTTTATCATGAGGCGCATCATAGCTTGCAGGCTTTATGCACTTCTGAAACAGAGGGGAC 1672  
QY 1796 TGGGGGTTCCCACTAGCACCCCACTTAACCTGATCAGCAGGAGCAGTGGTGGGAAAG 1855  
DB 1673 TGGGGGTTCCCACTAGCACCCCACTTAACCTGATCAGCAGGAGCAGTGGTGGGAAAG 1732  
QY 1856 CGGCCAGCTGTGCTTCTCTCCATGAGGACCACTCTCACTGCTTCCAGACAT 1915  
DB 1733 CGGCCAGCTGTGCTTCTCTCCATGAGGACCACTCTCACTGCTTCCAGACAT 1792  
QY 1916 TTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTCTTACAGTGG 1975



Db 520 CACAAGGGATGTTTCCACCACTATCACTGACAATGTACTGAATAGCAGCAGAGTCCAA 579  
Qy 421 GAGGCAATTGCGAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCAGCAGCAATCA 480  
Db 580 GAGGCAATTGCTGAGTGGCTGCGAATTTGAACCCGGATGGATCTGCCAGCAGAGTCC 639  
Qy 481 AAAGCGTTTAAACAAAGTGAAGAAAGAGCTTAAAGGATTTCTTCAAGAAATGGTTGCCACT 540  
Db 640 AAAGCCATCCAGAAAGTGAAGAAAGAGCCAGGAAGATCCTCCAGGAATGGTTGTACA 699  
Qy 541 GTCTCACCGGCAATGATTCAGACTGCTGGGTGGGTGCTGCTAAAACCTGTTCAACAGCTTC 600  
Db 700 GTCTCCCCCGGAGTATCAGGCTGACTGGCTGGGTGTTACTAAAGCTCTTCAACAGCTTC 759  
Qy 601 TTTTGGAAACATTCAAATTCACAAGGTCAACTTGAAGTGGTTAAAGCTGCACTGAGAGC 660  
Db 760 TTTCTGGAACATTCAGATTCACAAGGTCAACTTGAAGTGGTTAAAGCTGCACTGAGAGC 819  
Qy 661 AATTTGCCGCTTCTGTTTCTACCAAGTTCATAGATCCCATATTGACTATCTGTGCTCACT 720  
Db 820 AATCTGCCGCTCTGTTTCTGCCGCTGCACAGATCCCATCGACTACCTGCTGCTCACC 879  
Qy 721 TTCAATCTCTTTGCCATTAACCAAGCAGCAATACATTGCTTTCAGGCAATATCTCAAC 780  
Db 880 TTTCAATCTCTTTGCCATTAACCAAGTCCATACATCGCTCGGGCAACAACTCTCAAC 939  
Qy 781 ATCCCAATCTTCACTAGTACCTTGAATCCATAGCTTGGGGCTTCTTCAATCGAAGGCTC 840  
Db 940 ATCCCAATCTTCACTAGTACCTTGAATCCATAGCTTGGGGCTTCTTCAATCGAAGGCTC 999  
Qy 841 GATGAAACACCAAGTGAAGAAAGATGTTCTCTATAGAGCTTGTCTCATAGGGCATATA 900  
Db 1000 GACGAAACTCCAGTGAAGAAAGATGTTCTCTATAGAGCTTGTCTCATAGGGCATATA 1059  
Qy 901 GTTGAATTTACTCGAAGCAGCAATTTCTGGAGATCTTCTGGAAGGCAACGTTCTAGG 960  
Db 1060 GTTGAATCTCTCGAAGCAGCAATTTCTGGAGATCTTCTGGAAGGCAACCGCTCCCGC 1119  
Qy 961 AGTGGAAACCTTTGTGCTCGGCAGGACTTTTGTCTGATGTTGTGTAGATCTCTGCT 1020  
Db 1120 AGTGGCAAGACCTTCTGTGCTCGGCAGGACTTTTGTCTGATGTTGTGTAGATCTCTGCT 1179  
Qy 1021 ACCAATGTCTATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATCGCATATC 1080  
Db 1180 TCCAAACCATCTCTGATCTCTGTTCTATCCCTGTGGGATCTCTGATGATCGATATC 1239  
Qy 1081 GAAGGTCACTACAAATGGTGAACAACTGGGCAAACTTAAGAAAGTGAAGCCTGTGGAGT 1140  
Db 1240 GAAGGTCACTACAAATGGTGAACAACTGGGCAAACTTAAGAAAGTGAAGCCTGTGGAGT 1299  
Qy 1141 GTAGCAAGAGGTGTTATTAAGATGTTACGAAAGAACTATGTTGTGTCTCGAGTGGATTT 1200  
Db 1300 GTGGCAAGAGGTGTTATTAAGATGTTACGAAAGAACTATGTTGTGTCTCGAGTGGATTT 1359  
Qy 1201 GCACAGCAATTTTCTTAAAGAAATTTAGAAAGCAAAAGTCAAGAAACCGGTGTCTGCT 1260  
Db 1360 GCACAGCAATTTTCTTAAAGAAATTTAGAAAGCAAAAGTCAAGAAACCGGTGTCTGCT 1419  
Qy 1261 CTACTTTCCCTGGAGCAAGGTTGTTTCCAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320  
Db 1420 CCCCTCTCTTTGGAGCAAGCACTGTTTACAGCAATCTTCTTCAAGACCTGTATGCTGCT 1479  
Qy 1321 GCTGATGAAGGTAGACACAGTCCATTAATGATCCAGAAATGCAACAGATGATCCCTTA 1380  
Db 1480 GCTGCCGAAATGAAGACATGTCAGTAATGATGTCGAGAAACCGGCAAGCAAGGCTTC 1539  
Qy 1381 CGAAGGAGGTGATTCGAAATCTGGCTGAGCATATTTCTATTCTACTGCTAGCAAGTCTGT 1440  
Db 1540 CGAAGGAGGTGATTCGAAATCTGGCTGAGCATATTTCTATTCTACTGCTAGCAAGTCTGT 1599  
Qy 1441 GCCATTTATGTCCACACATTTGTGGCTTGTGCTCTCTCTACAGACACAGGCGGGAAT 1500

Db 1600 GCTATCATGTCCACCCACATTTGTGGCTTGTGCTCTCTACAGACACAGGCGGGAATC 1659  
Qy 1501 GATCTCTCCACATTTGTGGAAGACTTCTTTGTGATGAAGAGAAAGTCTCTGCTCTGTGAT 1560  
Db 1660 CACCTCTCCACGCTGTGGAAGACTTCTTTGTGATGAAGAGAAAGTCTCTAGCTCGGAT 1719  
Qy 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAGATGTAGTAATCATGCCCCATACAGTCTGTG 1620  
Db 1720 TTTGACCTGGGTTCTCAGGAAATTCAGAAGATGTAGTCATCATGCTATTTCAGCTTCTG 1779  
Qy 1621 GGAATTTGTGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1680  
Db 1780 GGAATTTGTGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1839  
Qy 1681 ACAATGTCTCCATCAGTCTTCCAACTTCTACAGCAATGGGTACTTCTCATGTCTTT 1740  
Db 1840 ACAATGTCTCCATCAGTCTTCCAACTTCTACAGCAATGGGTACTTCTCATGTCTTT 1899  
Qy 1741 ATCATGGAGGCTCATCATAGCTTTGAGCCCTTTATGAGCTTCTGAACAAGAGGAGCTGGG 1800  
Db 1900 ATCATGGAGGCTCATCATAGCTTTGAGCCATTTATGAGTCCAGAAATGAAGAGGCTTCCGGA 1959  
Qy 1801 GGTCCCATAGACACCCACCTTAACCTGATCAGCAGGAGGAGCTGGTGGGAGGCGGCC 1860  
Db 1960 GGTCTGCGGAGGCTTTGGCAACCTGATCAGCCAGGAGCAGTGGTGGGAGGCGGCC 2019  
Qy 1861 AGCCTGTGCTACTTCTCTCCAAATGAAGCACCCTCTCACTGCTTGCCTGCCAGACATTTTAC 1920  
Db 2020 AGCCTGTGCTACTTCTCTCTAATGAAGTACCATTTCTCTGCCCTGCCAGACATTTTAC 2079  
Qy 1921 CAAAGTCTGCCATGAACAGTAGGAAAGTTTATCCAGTATGGCAATCTTTACAGTGGCAGAG 1980  
Db 2080 CAGGTTTGTCAAGACAGTAGGAAAGTTTATCCAGTACGGAATTTCTCACAGTGGCAGAG 2139  
Qy 1981 CAGATGACAGGAAAGATATCAGTCTTAGTCTTGTGAGCAGCAGTGGGAGAAAGCTT 2040  
Db 2140 CAAAGTGAACAGGAAAGATGTCAGTCTTGGCCCTTGACAGAGCAGTGGGAAAGAGCTT 2199  
Qy 2041 CCAGAACCTTTCTCTTGGAGAGTGAAGAAAGATGAAGAGAGTGAAGAGTCTTTGGGAGGAA 2100  
Db 2200 CCGAGACCTCTGACTGAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAGTCTTTGGTGAAGAG 2259  
Qy 2101 CAGCAGATTTGCTACTGAAAGTGAAGCCAAATCCAAAGGAGCACCAGCAGTTTATCACCTTC 2160  
Db 2260 CAGCAGTGTGCTACTGAAAGTGAAGCCAAAGGAGCACCAGCAATTTATCACCTTT 2319  
Qy 2161 TTACAGAGACTCTTGTGGGCTTTGTGAGGAGCTTACAGCTCTGTGCTGCTTGTGTTAC 2220  
Db 2320 CTGCAAGGCTTCTGGGGCCCTGTAGAAAGCTTACAGCTCTGTGCTGCTTGTGCTAC 2379  
Qy 2221 AACTTCAGTGGTCTTCTCCAGAACTGAGTATCTGCAAAAGTTCACAAATACCTTAATA 2280  
Db 2380 ACTTTCGGGGCCAGTCCCGAGTCTGAGTACCTGCAAGAGTGGCAGAGTACCTTCTC 2439  
Qy 2281 ACCAGAAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGGCAGCATATTTGTCTTGTGAAG 2340  
Db 2440 ACCAGAACGAGAGGAAAGTGGCGGTGTACGCTGAGAGTGGCAGCATATCTGTCTGTGAAG 2499  
Qy 2341 AATGTGTGAAATGTTTAAAGATATTTGGGTTTTCAGGAGACCAAAACAAAGAGAGTG 2400  
Db 2500 AATGTGTGAAATGTTTAAAGATATTTGGGTTTTCAGGAGACCAAAACAAAGAGAGTG 2559  
Qy 2401 TCTGTTTTAGACTCAGCAGACTTTTCTTACCTCAATGCAACCGCAAAACCTTCTAGAA 2460  
Db 2560 TCTGTTTTAGACTCAGCAGACTTTTCTTACCTCAATGCAACCGCAAAACCTTCTAGAA 2619  
Qy 2461 TATATTTCTGAGTTTGTGGTGTGCTG 2484  
Db 2620 TACATTTCTGAGCTTCTGTGTGCTG 2643



AB08415 standard; cDNA; 469 BP.  
ABA08415;  
11-JAN-2002 (first entry)  
Human sn-glycerol-3-P acyltransferase homologue cDNA, SEQ ID NO:191.  
Human; cytokine; cell proliferation; cell differentiation; growth factor;  
haematopoiesis regulation; tissue growth; immunomodulator; activin;  
inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
chronic inflammatory condition; proliferative retinopathy;  
atherosclerosis; coronary heart disease; arterial ischaemia;  
bone disorder; osteoporosis; vascular growth disorder;  
tissue regeneration; wound healing; infection; immune disorder;  
cell culture; drug screening; gene therapy; antiinflammatory;  
antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
antifungal; vulnery; antiulcer; ss.  
Homo sapiens.  
WO200157188-A2.  
09-AUG-2001.  
05-FEB-2001; 2001WO-US03800.  
03-FEB-2000; 2000US-0496914.  
27-APR-2000; 2000US-0560875.  
(HYSE-) HYSEQ INC.  
Tang YT, Liu C, Drmanac RT;  
WPI; 2001-457740/49.  
P-PSDB; ABB11171.  
Human proteins and DNA encoding sequences useful for preventing,  
treating or ameliorating a medical condition in a mammalian subject  
e.g. arthritis and cancer -  
Claim 1; Page 405; 1963pp; English.  
Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
invention also relates to vectors and recombinant host cells comprising a  
nucleotide of the invention, methods of producing the novel polypeptides,  
antibodies against the polypeptides, methods of detecting the nucleotides  
or polypeptides in a sample, and methods of identifying compounds which  
bind to polypeptides of the invention. Although novel, many of the  
polypeptides of the invention have homology to known proteins, thereby  
giving an insight into their probable biological activities, and hence  
potential therapeutic applications. The polypeptides of the invention may  
have various activities, including cytokine, cell proliferation or cell  
differentiation activities; stem cell growth factor activity;  
haematopoiesis regulatory activity; tissue growth activity;  
immunomodulatory activity; activin- or inhibin-related activities;  
chemotactic or chemokinetic activities; haemostatic, thrombotic or  
thrombolytic activities; receptor or ligand activities; or may be  
involved in oncogenesis; cancer cell proliferation or metastasis.  
Depending on their biological activities, polypeptides and nucleotides of  
the invention are useful for preventing, treating or ameliorating medical  
conditions, e.g., by protein or gene therapy. Such conditions include  
cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
proliferative retinopathy, atherosclerosis, coronary heart disease,  
arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
vascular growth. Polypeptides involved with tissue regeneration and  
repair (or nucleic acids encoding them) may be used to promote wound  
healing (e.g., of burns, incisions and ulcers), while those with

immunomodulatory activities may be used in the treatment of viral,  
bacterial and fungal infections in addition to immune disorders.  
Polypeptides with growth factor activity may be used in cell cultures to  
promote cell growth. For example, such polypeptides may be used to  
manipulate stem cells in culture to give rise to neuroepithelial cells  
that can be used to augment or replace cells damaged by illness,  
autoimmune disease or accidental damage. The polypeptides and nucleotides  
may also be used in the diagnosis of the above conditions, and in drug  
screening techniques. The present sequence represents a cDNA encoding a  
novel human polypeptide of the invention.  
Sequence 469 BP; 142 A; 98 C; 109 G; 120 T; 0 other;  
Query Match 17.6%; Score 437; DB 22; Length 469;  
Best Local Similarity 100.0%; Pred. No. 9.2e-126;  
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 264 GCGAATGTTATTATATATCAATGAATCAACAGACACCGCGGATGGCTTGAAGACG 323  
Db 33 GCGAATGTTATTATATATCAATGAATCAACAGACACCGCGGATGGCTTGAAGACG 92  
Qy 324 CCTTTCTAGCTCTCTTTTATTCAAGAGCGAGATGTCATAAGGGCATGTTTGCACCA 383  
Db 93 CCTTTCTAGCTCTCTTTTATTCAAGAGCGAGATGTCATAAGGGCATGTTTGCACCA 152  
Qy 384 TGTGACTGAAATGCTGCTGAACAGCAGTAGAGTACAGAGGCAATTCAGAAATGGCTGC 443  
Db 153 TGTGACTGAAATGCTGCTGAACAGCAGTAGAGTACAGAGGCAATTCAGAAATGGCTGC 212  
Qy 444 TGAATTAACCCCTGATGGTTCTGCCAGAGCAATCAAAAGCCGTTAACAAAGTGA AAAA 503  
Db 213 TGAATTAACCCCTGATGGTTCTGCCAGAGCAATCAAAAGCCGTTAACAAAGTGA AAAA 272  
Qy 504 GAAGCTTAAAGGATCTTCAAGAAATGGTGGCCACTGTCTCACCGGCAATGATCAGACT 563  
Db 273 GAAGCTTAAAGGATCTTCAAGAAATGGTGGCCACTGTCTCACCGGCAATGATCAGACT 332  
Qy 564 GACTGGGTGGTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 623  
Db 333 GACTGGGTGGTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 392  
Qy 624 AGGTCAACTTGAGATGTTAAAGCTGCAACTGAGAGCAATTCGCGCTTCTGTTTCTACC 683  
Db 393 AGGTCAACTTGAGATGTTAAAGCTGCAACTGAGAGCAATTCGCGCTTCTGTTTCTACC 452  
Qy 684 AGTTCATAGATCCCATTA 700  
Db 453 AGTTCATAGATCCCATTA 469  
RESULT 10  
ABX42219  
ID ABX42219 standard; cDNA; 429 BP.  
XX  
AC ABX42219;  
XX  
DT 20-FEB-2003 (first entry)  
XX  
DE Bovine EST associated with lactation/muscle/fat deposition #7384.  
XX  
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
KW muscle deposition; fat deposition; genome mapping; gene identification;  
KW gene analysis; cattle breeding.  
XX  
OS Bos Taurus.  
XX  
XX US2002137139-A1.  
XX  
PN 26-SEP-2002.  
PD  
XX 24-SEP-2001; 2001US-0960352.  
PF  
XX 12-JAN-1999; 99US-115707P.  
PR







ID	ABX47921 standard; cDNA; 303 BP.	
XX	AC	ABX47921;
XX	XX	21-FEB-2003 (first entry)
DT	DT	Bovine EST associated with lactation/muscle/fat deposition #13086.
XX	XX	Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
XX	XX	muscle deposition; fat deposition; genome mapping; gene identification;
XX	XX	gene analysis; cattle breeding.
XX	XX	Bos Taurus.
OS	XX	US2002137139-A1.
XX	XX	26-SEP-2002.
XX	XX	24-SEP-2001; 2001US-0960352.
XX	XX	12-JAN-1999; 99US-115707P.
PR	PR	11-JAN-2000; 2000US-0480902.
XX	XX	(BYAT/) BYATT J C.
PA	PA	(MATH/) MATHIALAGAN N.
PA	PA	(TAON/) TAO N.
PA	PA	(WARR/) WARREN W C.
XX	XX	Byatt JC, Mathialagan N, Tao N, Warren WC;
PI	PI	WPI; 2003-110599/10.
DR	DR	New nucleic acid associated with lactation, and muscle and fat
XX	XX	deposition, useful for genome mapping, gene identification and
PT	PT	analysis, cattle breeding, or for genetically improving cattle
PT	PT	Claim 2; SEQ ID No 13086; 245pp; English.
XX	XX	The invention relates to a purified nucleic acid molecule associated with
CC	CC	lactation or muscle and fat deposition (designated LMFD), derived
CC	CC	from cattle, and the LMFD nucleic acid can specifically hybridise to a
CC	CC	second nucleic acid molecule comprising any of 15112 nucleotide
CC	CC	sequences, appearing as ABX34836-ABX49947, or complements of them.
CC	CC	Also included are; (1) a transformed cell having a nucleic acid
CC	CC	comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC	CC	translated sequence that functions in the cell to cause termination of
CC	CC	transcription and addition of polyadenylated ribonucleotides to a 3' end
CC	CC	of the mRNA molecule; and (2) determining a level or pattern of a
CC	CC	nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC	CC	complement or fragment) with a complementary nucleic acid molecule
CC	CC	obtained from the bovine cell or tissue, where hybridisation between the
CC	CC	marker nucleic acid and the complementary nucleic acid permits the
CC	CC	detection of the molecule; and (b) detecting the level or pattern of the
CC	CC	complementary nucleic acid, where the detection of the complementary
CC	CC	nucleic acid is predictive of the level or pattern of the molecule.
CC	CC	The LMFD nucleic acid is used for determining a level or pattern
CC	CC	of a molecule in a bovine cell or tissue. It is useful for genome
CC	CC	mapping, gene identification and analysis, cattle breeding, preparation
CC	CC	of constructs for use in cattle gene expression, or for genetically
CC	CC	improving cattle. The present sequence is one of the 15112 bovine
CC	CC	LMFD EST (expressed sequence tag) nucleic acids.
CC	CC	Note: The present sequence was not shown in the specification but
CC	CC	was obtained in electronic format from the USPTO web site:
CC	CC	seqdata.uspto.gov/sequence.html?DocID=20020137139.
XX	XX	Sequence 303 BP; 66 A; 90 C; 64 G; 83 T; 0 other;
XX	XX	Query Match 9.4%; Score 233; DB 25; Length 303;
XX	XX	Best Local Similarity 87.9%; Pred. No. 5.3e-62;
XX	XX	Matches 254; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY	QY	1907 GCCAGACATTTTACCAAGTCTGCCATGAACAGTAGGAAGTTTATCCAGTATGGCATTC 2466

CC obtained from the bovine cell or tissue, where hybridisation between the  
CC marker nucleic acid and the complementary nucleic acid permits the  
CC detection of the molecule; and (b) detecting the level or pattern of the  
CC complementary nucleic acid, where the detection of the complementary  
CC nucleic acid is predictive of the level or pattern of the molecule.  
CC The LMFD nucleic acid is used for determining a level or pattern  
CC of a molecule in a bovine cell or tissue. It is useful for genome  
CC mapping, gene identification and analysis, cattle breeding, preparation  
CC of constructs for use in cattle gene expression, or for genetically  
CC improving cattle. The present sequence is one of the 15112 bovine  
CC LMFD EST (expressed sequence tag) nucleic acids.  
CC Note: The present sequence was not shown in the specification but  
CC was obtained in electronic format from the USPTO web site:  
CC seqdata.uspto.gov/sequence.html?DocID=20020137139.  
XX  
SQ Sequence 264 BP; 75 A; 60 C; 56 G; 73 T; 0 other;

Query Match 8.4%; Score 209.6; DB 25; Length 264;  
Best Local Similarity 90.4%; Pred. No. 1e-54;  
Matches 235; Conservative 0; Mismatches 24; Indels 1; Gaps 1;  
Oy 1468 TGCCTGCTCCTACAGACACAGGAGGAATTGATCTCCACATTTGGTGAAGACTTC 1527  
Db 2 TGTCTGCTCCTACAGACACAGGAGGAATTGATCTCCACATTTGGTGAAGACTTC 61  
Oy 1528 TTTGTGATGAAGAGGAAGTCTGCTCTGTGATTTTGACCTGGGTTCTCAGGAATTC 1587  
Db 62 TTTGTGATGAAGAGGAAGTCTGCTCTGTGATTTTGACCTGGGTTCTCAGGAATTC 121  
Oy 1588 G-AGATGTAGTAATGCATGCATACAGTCTGCGGAATTTGTGCACAATCACCAC 1646  
Db 122 GAAAGTGTAGTCATGCATGCCAAACAGCTCCTGGGAATTTGTATCAAAATCACCAC 181  
Oy 1647 TAGCAGGAACGATGAGTTTTTTTATCACCCCGACCAACTGTCCTCAGTCTTGGAACT 1706  
Db 182 CAGCAAGATGATGAGTTTTTTTATTTACTCTAGCAACAATATCCCATCAGTCTTGGAACT 241  
Oy 1707 CAATCTTACAGCAATGGGG 1726  
Db 242 CAATCTTCAAAACAAATGGG 261

RESULT 15  
ABX42469  
ID ABX42469 standard; cDNA; 387 BP.  
AC ABX42469;  
XX  
DT 20-FEB-2003 (first entry)  
DE Bovine EST associated with lactation/muscle/fat deposition #7634.  
XX  
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
KW muscle deposition; fat deposition; genome mapping; gene identification;  
KW gene analysis; cattle breeding.  
XX  
OS Bos Taurus.  
XX  
PN US2002137139-A1.  
PD  
XX  
PD 26-SEP-2002.  
XX  
PD 24-SEP-2001; 2001US-0960352.  
PR 12-JAN-1999; 99US-115707P.  
PR 11-JAN-2000; 2000US-0480902.  
XX  
PA (BYAT/) BYATT J C.  
PA (MATH/) MATHIALAGAN N.  
PA (TAON/) TAO N.  
PA (WARR/) WARREN W C.  
XX  
PI Byatt JC, Mathialagan N, Tao N, Warren WC;

XX  
DR WPI; 2003-110599/10.  
XX  
PT New nucleic acid associated with lactation, and muscle and fat  
PT deposition, useful for genome mapping, gene identification and  
PT analysis, cattle breeding, or for genetically improving cattle  
XX  
PS Claim 2; SEQ ID No 7634; 245pp; English.  
XX  
CC The invention relates to a purified nucleic acid molecule associated with  
CC lactation or muscle and fat deposition (designated LMFD), derived  
CC from cattle, and the LMFD nucleic acid can specifically hybridise to a  
CC second nucleic acid molecule comprising any of 15112 nucleotide  
CC sequences, appearing as ABX34836-ABX49947, or complements of them.  
CC Also included are: (1) a transformed cell having a nucleic acid  
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-  
CC translated sequence that functions in the cell to cause termination of  
CC transcription and addition of polyadenylated ribonucleotides to a 3' end  
CC of the mRNA molecule; and (2) determining a level or pattern of a  
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker  
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its  
CC complement or fragment) with a complementary nucleic acid molecule  
CC obtained from the bovine cell or tissue, where hybridisation between the  
CC marker nucleic acid and the complementary nucleic acid permits the  
CC detection of the molecule; and (b) detecting the level or pattern of the  
CC complementary nucleic acid, where the detection of the complementary  
CC nucleic acid is predictive of the level or pattern of the molecule.  
CC The LMFD nucleic acid is used for determining a level or pattern  
CC of a molecule in a bovine cell or tissue. It is useful for genome  
CC mapping, gene identification and analysis, cattle breeding, preparation  
CC of constructs for use in cattle gene expression, or for genetically  
CC improving cattle. The present sequence is one of the 15112 bovine  
CC LMFD EST (expressed sequence tag) nucleic acids.  
CC Note: The present sequence was not shown in the specification but  
CC was obtained in electronic format from the USPTO web site:  
CC seqdata.uspto.gov/sequence.html?DocID=20020137139.  
XX  
SQ Sequence 387 BP; 87 A; 97 C; 113 G; 90 T; 0 other;

Query Match 6.9%; Score 172.6; DB 25; Length 387;  
Best Local Similarity 88.6%; Pred. No. 5.1e-43;  
Matches 187; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
Oy 2274 CCTAATAACCAAGACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCCACATATTGTCT 2333  
Db 1 CCTGATCAGGAGAACCGAGAGAGGGTTGCCGTGACGTGAGAGCGCCATTACTGTCT 60  
Oy 2334 TGTGAAGAATGCTGTGAAAATGTTTAAAGGATATTGGGGTTTTCAAGGAGACCAACAAAA 2393  
Db 61 CGTGAAGAATGCTGTGAAGACGTTTAAAGGATATCGGGTTTTTCAAGAGACCAACAAAA 120  
Oy 2394 GAGAGTCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAATGCAACCGACAAAAACT 2453  
Db 121 GAGAGTCTGTTTTAGAACTCAGCAACACTTTTCTACCTCAATGCAACCGGAAAAACT 180  
Oy 2454 TCTAGAATATATTCTGAGTTTTTGTGGTGCTG 2484  
Db 181 CCTGAATATATTCTGAGTTTTTGTGGTGCTG 211

Search completed: January 10, 2004, 00:26:35  
Job time : 430.997 secs

**THIS PAGE BLANK (USPTO)**